RHV

Revista de Humanidades de Valparaíso

Revista internacional de filosofía

Año 7, 2019, 1er Semestre, No 13

Humanities Journal of Valparaiso

An International Journal of Philosophy

No 13 (2019)

Universidad de Valparaíso

Facultad de Humanidades

Instituto de Filosofía

Revista de Humanidades de Valparaíso (RHV)
Humanities Journal of Valparaiso

eISSN 0719-4242 – https://revistas.uv.cl/index.php/RHV/
CDD: 090

No 13 (2019) - DOI: https://doi.org/10.22370/rhv2019iss13

Contacto / Contact: rhv.editores@gmail.com

Auspicio y Patrocinio / Sponsorship: Universidad de Valparaíso

Nota del editor

La *Revista de Humanidades de Valparaíso* (RHV) es editada por el Instituto de Filosofía de la Facultad de Humanidades de la Universidad de Valparaíso desde el año 2013. Su periodicidad de publicación es bianual de artículos inéditos y reseñas bibliográficas del área de la filosofía. La RHV publica en cuatro idiomas (castellano, portugués, inglés y francés), no se suscribe a ninguna doctrina particular y está abierta a artículos de diferentes perspectivas filosóficas y con un alcance internacional.

Editor's Note

The *Humanities Journal of Valparaiso* (RHV, for its acronym in Spanish) is edited by the Institute of Philosophy of the Faculty of Humanities of the University of Valparaiso since 2013. Its periodicity is biannual for unpublished works in the field of philosophy. The RHV published in four languages, Spanish, Portuguese, English and French; and does not subscribe to any particular doctrine and is open to articles from different philosophical perspectives and with an international scope.

TABLA DE CONTENIDOS

TABLE OF CONTENTS

Revista de Humanidades de Valparaíso
Año 7, 2019, 1er semestre, No 13, págs. 6-17
eISSN 0719-4242

Humanities Journal of Valparaiso
No 13 (2019): 6-17

A influência do Tractatus no critério positivista de significado

The Influence of Tractatus in the Positivist Criterion of Meaning

Renata Maria Santos Arruda
Universidade Federal de Goiás, Brasil
renataarrudaprof@gmail.com

DOI: https://doi.org/10.22370/rhv2019iss13pp6-17
Recibido: 27/05/2019. Aceptado: 14/08/2019

Resumo

Uma das motivações para as pesquisas realizadas pelos membros do Círculo de Viena a respeito dos fundamentos da linguagem científica se encontra na obra de Wittgenstein "*Tractatus Logico-Philophicus*". Embora haja divergências sobre a legitimidade dessa influência, o livro foi, com efeito, tomado como motivação teórica para a estruturação da linguagem científica, desenvolvida pelos empiristas lógicos. Este artigo apresenta as teorias desenvolvidas pelos membros do Círculo de Viena ressaltando os elementos presentes no *Tractatus* que foram tomados, em grande parte, como influência sobre o "critério de significado" de uma proposição filosófica adotado pelos empiristas lógicos na tentativa de se estabelecer uma fundamentação para o conhecimento.

Palavras-chave: Círculo de Viena, Wittgenstein, estrutura lógica, correspondência, verificacionismo.

Abstract

One of the motivations for the researches leaded by the members of Vienna Circle with regard to the foundations of scientific language is found on Wittgenstein's "*Tractatus Logico-Philophicus*". Even though there are divergences about the legitimacy of this influence, the book was, in effect, taken as a theoretical motivation for the structuring of scientific language, developed by logical empiricists. This paper will present the theories developed

by members of Vienna Circle emphasizing the elements presents in *Tractatus* that were taken, largely, as influence over the "criterion of meaning" of a philosophical proposition adopted by logical empiricists in an attempt to set a foundation for knowledge.

Keywords: Vienna Circle, Wittgenstein, logical structure, correspondence, verificationism.

1. Introdução

Há exatos 90 anos, em agosto de 1929, o movimento filosófico engendrado pelo Círculo de Viena demarcava suas perspectivas filosóficas por meio da publicação de seu "Manifesto", marcando a consolidação de uma nova abordagem da ciência na história da Filosofia. Meu objetivo nesse artigo é celebrar esse momento histórico pelo resgate de alguns elementos desse movimento relacionados aos seus posicionamentos teóricos iniciais; mais especificamente, a controversa influência de algumas ideias do Tractatus na forma em que os pensadores do Círculo viam a relação entre a realidade e a linguagem.

Com a nomeação do físico Moritz Schlick para ocupar a cadeira de Filosofia e de História das Ciências Indutivas na Universidade de Viena, em 1929, formou-se um grupo que, frente aos avanços tecnológicos e científicos da época, propôs-se a discutir os rumos da ciência e o papel da filosofia nesse novo contexto. Entre os simpatizantes desse movimento estavam A. J. Ayer, Carl G. Hempel, Hans Reichenbach, entre outros, sendo que seus membros principais eram Kurt Gödel, Hans Hahn, Otto Neurath, Viktor Kraft e Rudolf Carnap. Enquanto corrente filosófica, a orientação teórica daqueles pensadores foi caracterizada como Neopositivismo ou Positivismo Lógico. Essa corrente tinha como objetivo não só aproximar ciência e filosofia por meio do empirismo e separar esta da metafísica, mas também renovar os problemas filosóficos por meio do aspecto formal do conhecimento (Hahn et al.1986, 9). Sob esse título, as pesquisas dos seus membros buscavam uma interpretação do conhecimento em termos lingüísticos; a investigação empírica, que atribui validade ao conhecimento resultado das experiências, do contato com o mundo externo, seria pautada por um critério lógico que determinaria a validade do conhecimento adquirido nesse tipo de investigação.

Por meio da estruturação da linguagem lógica e do reducionismo aos termos observacionais, que se conectam mais diretamente à experiência por referência direta aos objetos físicos, os pensadores do movimento positivista também se mobilizaram na tentativa de unificação da linguagem da ciência em uma linguagem lógica, científica, objetiva, em detrimento de uma linguagem natural, onde seriam possíveis significações subjetivas. Com um vocabulário definido universalmente seria possível que os termos de toda a ciência se referissem a um significado determinado, onde os dados empíricos unificassem diferentes teorias, permitindo o intercâmbio inequívoco de idéias entre os pesquisadores sobre uma base empírica comum, o que viria a contribuir para o progresso e desenvolvimento da ciência (Hahn et al. 1986, 12). Com o intuito de estabelecer uma unificação da linguagem

da ciência, os filósofos do positivismo lógico procuraram estruturar uma linguagem com base em certos conceitos desenvolvidos no *Tractatus* de Wittgenstein, como veremos a seguir.

2. As teses do Tractatus relevantes para o positivismo lógico

Toda a análise acerca da estrutura lógica no *Tractatus* é desenvolvida por meio do traço fundamental da filosofia, que consiste, para Wittgenstein, na propriedade da elucidação de conceitos, de clarificação de idéias. A filosofia permite lançar luz sobre as relações implícitas nas proposições da linguagem e, nesse sentido, revela a natureza e o sentido dessas proposições. Enquanto expõe a estrutura lógica, a filosofia delimita a expressão do pensamento e da linguagem e revela, conseqüentemente, os limites do pensamento e as possibilidades da linguagem.

Um dos pressupostos das ideias desenvolvidas no *Tractatus* acerca da linguagem é que o sentido das proposições compostas é compreendido "de baixo para cima", o que exige que se deva conhecer inicialmente o significado das palavras. Esse processo se inicia com a compreensão do significado das palavras, que determinam o significado das proposições elementares. Por outro lado, devem-se pressupor também certas regras de ligação entre essas palavras, regras que estabeleçam os sentidos das diferentes frases através da gramática de uma determinada língua. O caráter do composicionalismo do *Tractatus* é descrito com bastante clareza no aforismo 4.024: "Entende-se uma proposição caso se entendam suas partes constituintes". Em outros termos, Wittgenstein indagaria: como entender o significado de uma proposição nunca antes ouvida senão pela compreensão do significado de suas palavras? Sua semântica composicionalista responde de maneira adequada à essa questão, e demonstra que com as mesmas palavras e expressões novos sentidos podem ser comunicados. Dessa forma, as palavras formam uma espécie de quebra-cabeça, e as regras gramaticais estabelecem determinados encaixes que permitem diferentes tipos de peças, as quais podem ser permutadas para oferecerem diferentes aspectos acerca do todo.

A conexão dessa estrutura da linguagem com a realidade é discutida no *Tractatus* pela compreensão de que a linguagem reflete, em um sentido imagético, estrutural, a ordem do mundo, à medida que a linguagem fale sobre a ocorrência ou não na *realidade* do fato ou do estado de coisas que ela expressa. Mas a verdade de uma sentença pode também ser tautológica, analítica, verdadeira apenas em virtude da sua forma lógica, e, assim, não diz nada sobre o mundo. Quando se trata de um fato expresso na linguagem, o modo como os elementos dos estados de coisas estão uns para os outros é representado pelas proposições elementares, que, conseqüentemente, são formadas pelos nomes simples, os quais representam objetos simples. Logo, pela associação das proposições elementares em uma função de verdade, o valor de verdade de uma proposição complexa depende, em última análise, de que a denominação dos objetos pelos nomes simples seja a adequada, isto é, que aquilo que os nomes simples nomeiam represente os próprios objetos a que eles se referem.

A análise lógica de uma proposição complexa revela a sua formação a partir de proposições elementares, e cada uma dessas proposições afigura um estado de coisas. A combinação possível – lógica – desses estados de coisas é representada pelas proposições complexas. Em outras palavras, a verdade de uma proposição complexa depende da operação de verdade das proposições elementares na medida em que estas representem um estado de coisas existente, isto é, na medida em que forem um fato. Assim, a função de verdade formada pela combinação das proposições elementares para a definição da verdade da proposição complexa, permite, enfim, afigurar a disposição dos objetos do mundo, como eles estão uns para os outros, na proposição complexa.

Wittgenstein traça um paralelo entre a linguagem e o mundo, onde a linguagem reflete a estrutura lógica do mundo, compondo sua forma da realidade. Na forma da realidade, a relação entre os objetos da realidade é comparada à relação entre os objetos da afiguração: os elementos da figuração devem estar relacionados como estão as coisas no mundo. A forma de afiguração, que é a forma lógica comum à figuração e ao afigurado, ao possibilitar as inter-relações entre os elementos afigurados, ou seja, a combinação de elementos lingüísticos representando um estado de coisas possível pode realizar uma descrição verossímil feita pela linguagem acerca do mundo. Nessa perspectiva, revela-se o aspecto correspondencionalista do *Tractatus* através da representação do mundo por meio da linguagem. É importante ressaltar que a estrutura lógica é representada não como um objeto do mundo, mas como uma relação existente entre os objetos afigurados. Pois se a forma lógica pudesse ser representada, seria equivalente aos elementos afiguráveis, e a cada representação lógica dos elementos afiguráveis decorrer-se-ia outra representação que, por sua vez, seria formalizada por uma outra representação, e assim por diante, configurando-se um regresso ao infinito. Ou seja, a forma lógica, que é comum à linguagem e à realidade, não pode ser afigurada, ela é mostrada nas relações entre os objetos e as proposições.

Desse modo, a análise filosófica realizada no *Tractatus* expõe a forma do pensamento, da linguagem e do mundo, e mostra como eles são permeados por uma única estrutura lógica, sua forma de afiguração. Não podem existir, de acordo com o que é defendido no *Tractatus*, diferentes formas de se pensar logicamente. A lógica existente nas relações entre os objetos, nos estados de coisas, no pensamento e na linguagem, é uma só. Não existem, portanto, diferentes formas de se pensar o mundo com sentido, isto é, não existe uma maneira ilógica de se pensar com sentido.

Além das proposições com sentido, os outros tipos de sentenças que podem existir são os contrassensos, tautologias, contradições e os enunciados da matemática. Nos contrassensos, ocorre uma combinação de nomes logicamente incompatíveis, formando uma pseudoproposição e ultrapassando os limites da razão, pois, apesar de respeitarem as regras gramaticais para a construção de uma sentença inteligível, estes enunciados não afiguram um estado de coisas, não falam nada sobre o mundo. A combinação que ocorre nesse tipo de enunciado não é autorizada pelos fatos, não possui correlato ontológico, enfim, não é dada pela possibilidade de fatos no mundo. Por outro lado, as proposições sem

sentido, que dizem respeito às tautologias, contradições e aos enunciados da matemática, e não representam um estado de coisas em absoluto. Um estado de coisas é representado por uma proposição elementar; a proposição elementar diz algo sobre o mundo: descreve o que as coisas são e como elas estão, isto é, tem um sentido que deve poder ser verdadeiro ou falso, e as tautologias ou contradições dizem respeito ao que é necessariamente verdadeiro e necessariamente falso, respectivamente. As proposições sem sentido, portanto, não descrevem nenhum estado de coisas.

Algumas teses de Wittgenstein, presentes no *Tractatus,* tais como a teoria da figuração e certas noções sobre a tabela de verdade, foram aceitas em um primeiro momento nas pesquisas do Círculo de Viena[1]. É importante notar que Wittgenstein não participava das reuniões do Círculo de Viena e não compartilhava diretamente das teses desenvolvidas por seus membros, que, apesar disso, consideravam as teorias do *Tractatus* como precursoras do positivismo lógico, ganhando uma posição de destaque para a fundamentação, de um modo específico, do "critério de significado". Mesmo permanecendo distante das teorias influenciadas por ele, os seguintes livros da fase "verificacionista" de Wittgenstein foram o resultado de discussões com alguns membros mais influentes do Círculo de Viena, onde o filósofo participa esclarecendo conceitos e algumas questões referentes ao *Tractatus*: do diálogo com Schlick e Waismann, publicou-se postumamente "*The Principles of Linguistic Philosophy*"[2] (1965) e "*Ludwig* e em 1967 *Wittgenstein und der Wiener Kreis*" (1996), sendo que este consiste em anotações feitas por Waismann dos comentários de Wittgenstein; das anotações de G. E. Moore publicou-se em 1964, posteriormente ao *Tractatus,* "*Philosophische Bemerkungen*"[3] (Moore *apud* Barbosa Filho 1981, 21).

3. O critério positivista de significado

Na medida em que descrevia a estruturação do pensamento, do mundo e da linguagem, as teses do *Tractatus* forneciam elementos para a constituição de certas teorias dos Positivistas Lógicos (Barbosa Filho 1981, 22). Estes tomavam como ponto de partida a pesquisa empírica, onde desenvolviam teorias correspondencionalistas sobre o conhecimento, justificando a verdade como correspondência à realidade empírica e assumindo um compromisso lógico com a coerência dos sistemas teóricos (Hempel 1935, 49). O discurso acerca do conhecimento deveria estar fundamento nessas duas bases, pois estaria

[1]Encontra-se também em textos posteriores ao *Tractatus* o desenvolvimento da noção de verificabilidade, na chamada fase verificacionista (Monk 1995, 21) ou fenomenalista (Barbosa Filho 1981, 21) de Wittgenstein.

[2]Segundo Monk (1995, 260), o livro em questão serviria como uma introdução às idéias do *Tractatus*, e fora composto por notas tomadas por Waissman e Schlick em reuniões com Wittgenstein, onde este discutia a respeito do *Tractatus*. O livro não foi publicado àquela época, pois Wittgenstein se negou, posteriormente, a colaborar com os positivistas.

[3]De acordo com Barbosa Filho (1981, 21), as anotações de Moore que deram origem ao livro supracitado se referem às aulas proferidas por Wittgenstein entre 1930 e 1933. Em Barbosa Filho (*ibíd.*, 31) estão relacionados outros nove importantes livros de Wittgenstein que dizem respeito à temática do Círculo de Viena.

comprometido com o mundo e com o pensamento, e, assim sendo, a sua plausibilidade estaria garantida de modo completo pelos meios considerados indicadores da verdade: a experiência e a lógica.

O empirismo lógico, na sua busca por uma interpretação do conhecimento em termos lingüísticos, tentou estabelecer uma fundamentação para o conhecimento acerca do mundo em uma base empírica, admitida como o princípio lógico ou temporal da construção dos saberes humanos[4], sistematizada por uma linguagem adequada (Schlick 1988a, 66). O conhecimento legítimo seria, nesses termos, representado pela ciência natural, e a linguagem que ordenaria os conceitos científicos transformava-se numa metodologia embasada por conceitos da lógica. A ciência da época apresentava grandes avanços: a Geometria e as teorias da Física, a partir do desenvolvimento das geometrias não-euclidianas e da teoria da relatividade, perdiam o terreno do absoluto antes construído por uma única forma de descrição dos fenômenos físicos e geométricos (Hahn et al. 1986, 15). A matemática se renovara com as idéias desenvolvidas no *Tractatus*, a partir das contribuições de Frege e Russell para a matemática, na tentativa de atribuir a esta uma base lógica (Kenny 1995, 46). Nesse contexto, as considerações dos positivistas lógicos acerca da fundamentação do conhecimento acompanhavam as transformações pelas quais passava o conhecimento científico, prático e teórico. O pensamento da época se transformava com a busca de um conhecimento sólido, onde se pudesse, mesmo que em princípio, vislumbrar respostas definitivas para as questões que sempre motivaram as investigações sobre o mundo que nos cerca. Diferentemente da metafísica, a que os Positivistas atribuíam o empreendimento de especulações inesgotáveis sobre problemas insolúveis e conceitos que transcendem a experiência, os membros do Círculo de Viena, ao se debruçarem sobre o mundo físico através da ciência, tinham a pretensão de traduzir, por meio da Filosofia, algumas certezas que o mundo pudesse oferecer (Hahn et al. 1986, 10).

A influência do *Tractatus* sobre o "critério de significado" de uma proposição, adotado pelos positivistas lógicos, parte de uma interpretação epistemológica dos conceitos de função de verdade e figuração presentes nesse livro (Barbosa Filho 1981, 22). Wittgenstein não estava interessado em definir ou em apontar para qualquer objeto do mundo físico e denominá-los a partir dos conceitos da sua teoria. Sua preocupação era puramente lógica, formal. O que está pressuposto no *Tractatus* é a existência dos objetos para uma afiguração completa do mundo, tratando-se apenas das relações lógicas desses objetos, isto é, da forma lógica como são constituídos e do sentido da proposição complexa a partir do valor de verdade das suas proposições elementares. O positivismo lógico se encarregou, a partir da estrutura lógica da linguagem sistematizada pelo *Tractatus,* de acrescentar um conteúdo empírico às proposições lógicas. Esse conteúdo seria constituído por objetos do mundo físico, os quais seriam nomeados pelos nomes simples e descritos pelas proposições elementares e complexas.

[4]Esta opinião é expressa de um modo geral pelos positivistas lógicos, apesar de não acordarem se o princípio seria temporal ou lógico. A formulação aceita por Schlick, por exemplo, é que o princípio do conhecimento possa ser, simultaneamente, temporal e lógico.

O critério que determina o significado de um enunciado, adotado pelos positivistas lógicos, pode ser resumido na seguinte afirmação de Schlick:

> Enunciar o sentido de uma frase equivale a estabelecer as normas segundo as quais a frase deve ser empregada, o que significa enunciar a maneira pela qual se pode constatar a sua verdade (ou a sua falsidade). O significado de uma proposição constitui o método da sua verificação. (Schlick 1988b, 85)

O critério de significado dos positivistas está baseado na idéia do *Tractatus* que diz respeito à função de verdade de uma proposição complexa em relação a uma proposição elementar (Wittgenstein 2001, 63). Dessa forma, as proposições elementares só se tornam verdadeiras, e, conseqüentemente tornam as proposições complexas verdadeiras, se os nomes simples estiverem substituindo, na proposição, os objetos que se propõem a afigurar[5]. Os nomes simples, por sua vez, estariam em conexão direta com a realidade, seriam fixados por relação direta com o mundo. Assim sugere o *Tractatus* no aforismo 4.21: "A proposição mais simples, a proposição elementar, assere a existência de um estado de coisas"; em 4.22: "A proposição elementar consiste em nomes. É uma vinculação, um encadeamento de nomes", e em 3.22: "O nome substitui, na proposição, o objeto". Portanto, caso se reduzisse o sentido de uma proposição complexa ao significado dos seus termos componentes, uma proposição com sentido só poderia ser aquela que estivesse dada na realidade.

Pressupondo uma estrutura lógica na qual os elementos do mundo material pudessem ser inseridos, a teoria dos positivistas implicaria, desse modo, a definição de um significado a partir de uma linguagem e de regras lógicas, metodologicamente estabelecidas, que permitiriam definir a validade dos raciocínios desenvolvidos, associados a um conteúdo empírico necessariamente vinculado às proposições. Assim, a fundamentação do conhecimento dependeria do critério de significado para o desenvolvimento de seus conceitos basilares, pois, de posse do significado dos termos pertinentes a uma teoria, poder-se-ia descrever adequadamente as relações dos objetos no mundo e a verificação dos enunciados científicos.

A análise lógica e sua aplicabilidade ao material empírico tornaram possível aos positivistas lógicos, por meio da elucidação dos termos componentes de uma proposição complexa, estruturarem o reducionismo de conceitos superiores mais complexos a conceitos inferiores mais básicos, mais próximos da experiência (Hahn et al. 1986, 12). Esta estruturação formaria um sistema de constituição dos conceitos, que os ordenaria de acordo com a referência destes aos objetos físicos, o que tornaria a ciência unificada. A filosofia, no contexto teórico do positivismo lógico, tornou-se mais do que nunca filosofia da ciência, cabendo a ela a tarefa de elucidar as proposições científicas que se referem diretamente ou indiretamente à experiência e estabelecer o método logicamente adequado para

[5]De acordo com Barbosa Filho (1981, 21), Wittgenstein estabeleceu, nas "*Philosophische Bemerkungen*", o seu próprio critério de significado, relacionando as proposições elementares com as proposições fenomenais.

a verificação dos enunciados científicos, restando à ciência verificar aqueles enunciados (Hahn et al. 1986, 18). Através de uma linguagem unificada entre ciência e filosofia, pretendia-se chegar também a uma explicação unitária do mundo (Hahn et al. 1986, 10). Portanto, a assim chamada "concepção científica do mundo" do Círculo de Viena buscou desenvolver-se a partir da clareza e rigor de conceitos, exatidão na linguagem e no simbolismo lógico, independentemente da origem e da orientação científica.

A linguagem, aplicada ao material empírico, se estabelecia, dessa forma, como uma metodologia para a clarificação dos conceitos da ciência, e os membros do Círculo de Viena reconheciam como legítima a tarefa da filosofia, descrita no *Tractatus*, como "atividade" de elucidação do conhecimento. Esta proposta é amplamente divulgada no "manifesto" escrito por membros do Círculo de Viena, mas especialmente no seguinte trecho:

> A concepção científica do mundo *desconhece enigmas insolúveis*. O esclarecimento dos problemas filosóficos tradicionais conduz a que eles sejam parcialmente desmascarados como pseudoproblemas e parcialmente transformados em problemas empíricos sendo assim submetidos ao juízo das ciências empíricas. A tarefa do trabalho filosófico consiste neste esclarecimento de problemas e enunciados, não, porém, em propor enunciados "filosóficos" próprios. (Hahn et al. 1986, 10)

Ao repudiarem a metafísica em virtude da ausência de referencial empírico de seus termos, os positivistas lógicos também foram considerados "anti-filósofos" por se voltarem exclusivamente para as questões pertinentes ao campo que antes cabia exclusivamente à ciência, utilizando a lógica como ferramenta de elucidação (Neurath 2003, 121). Os positivistas acreditavam que de um conjunto de enunciados mal formulados e equivocadamente interpretados poderia se chegar a problemas inexistentes, decorrentes apenas dos erros lógicos da má constituição desses enunciados, que resultariam na má compreensão dos mesmos. A lógica, como instrumento de análise, dissecaria a linguagem em sua estrutura e em suas conexões com os nomes simples, os quais, finalmente mediariam as proposições com a realidade e evitariam os danos de uma descrição equivocada da realidade.

Vinculada à ciência, a Filosofia desenvolvida pelos Positivistas Lógicos a partir da leitura que fizeram do *Tractatus*, tinha como função fundamental analisar as proposições científicas até os seus termos últimos, identificar os seus objetos e, finalmente, apresentar o sentido dessas proposições, vinculando-as à experiência. O *Tractatus*, ao sistematizar a linguagem e ao conectá-la ao mundo, delineou a estrutura pela qual os positivistas instituiriam convencionalmente os objetos pertencentes à base empírica nas diferentes teorias desenvolvidas por eles, entre elas o fisicalismo e o fenomenalismo. Sob essa perspectiva, qualquer raciocínio só poderia ser intitulado de "filosófico" se seus elementos se remetessem à experiência.

Consequentemente, o princípio verificacionista também se tornou um critério de demarcação entre ciência e não-ciência. Por reduzir qualquer enunciado significativo a um enunciado observacional, os enunciados que não referissem a um conteúdo empírico se-

riam identificados como isentos de significado. Seriam assim considerados, da mesma forma que Wittgenstein define no *Tractatus*, pseudoproposições ou contra-sensos. Portanto, no contexto do positivismo lógico, assim como no *Tractatus*, só poderiam haver dois tipos de enunciados aceitos no corpo da ciência: as proposições tautológicas (as expressões matemáticas ou da lógica), pois essas dizem respeito às leis do pensamento, e as proposições empíricas, que seriam proposições com sentido (Hahn et al. 1986, 10).

É importante destacar que o significado de um enunciado qualquer está subsumido a determinadas regras lingüísticas, previamente estabelecidas; regras que visem definir a referência dos termos e o sentido das proposições. Somente de posse de um significado definido semanticamente e sintaticamente, podemos atribuir às proposições, em seguida, um valor de verdade. Que um enunciado possua um significado não se quer dizer que ele seja verdadeiro, mas sim que se têm as condições necessárias para afirmá-lo verdadeiro ou falso. E essa afirmação, na presente abordagem, constitui na verificação do fato ou estado de coisas expressos por esse enunciado.

O método de verificação dos enunciados da ciência sofreu importantes alterações ao longo do desenvolvimento das idéias propostas pelos positivistas lógicos, e a mudança mais significativa foi o estabelecimento de um convencionalismo para a base empírica. Com isso, os elementos últimos de qualquer enunciado seriam definidos a partir das condições e necessidades de contato com a realidade da teoria à qual o enunciado estivesse subsumido (Hempel 1935, 58). Nesse sentido, não haveria mais um único "critério de significado" e nem mesmo um consenso acerca da possibilidade de uma unificação da ciência por meio de uma única base empírica, porém, não se alteraria a necessidade de se estabelecer os elementos fundamentais para uma compreensão geral de um corpo de enunciados relativos a uma teoria.

4. Conclusão

A teoria do conhecimento desenvolvida pelos membros do Círculo de Viena a partir de suas leituras do *Tractatus*, que os unia em torno de um critério de significado, baseava-se na estipulação de uma interpretação das proposições filosóficas, e em um método para a verificação dessas proposições. Não é casual, portanto, a escolha dos termos "empirismo lógico" pelos membros do Círculo de Viena para definir seu posicionamento epistemológico, fazendo referência à aplicação da ferramenta lógica ao empirismo em seu sentido moderno, o qual encontrava na experiência sensível a origem das nossas ideias e do nosso conhecimento. A lógica se tornaria então uma linguagem específica que estruturaria e articularia os dados dos sentidos. Por meio da análise lógica pretendia-se definir a verificabilidade das proposições, revelando as condições nas quais as proposições poderiam ser testadas empiricamente pelas ciências. A filosofia passaria a ser uma ferramenta para a elucidação do pensamento, dos conceitos; a análise lógica revelaria o sentido do enunciado, mostrando o que ele realmente quer dizer, e seu significado estaria relacionado à possibilidade de verificá-lo na experiência. Estes aspectos são retratados, na linguagem

do *Tractatus*, como uma teoria da figuração, onde a linguagem corresponde aos fatos da realidade, e ambas, linguagem e realidade, se afiguram uma à outra, desde suas partes atômicas até as mais complexas, e o critério de significado dos positivistas pretendia aplicar essa correspondência no âmbito dos enunciados científicos. Em virtude da preocupação de que a linguagem empregada na filosofia deva refletir a realidade, a metafísica, entendida como parte da filosofia que investiga as essências das coisas para além do dado concreto, é criticada pelos empiristas lógicos exatamente por não se referir à experiência empírica, ao mesmo tempo em que, por exemplo, usa termos abstratos como indicando entidades concretas.

Curiosamente, a própria concepção do método de verificação engendrava a mudança pela qual passaria o critério de significado. Por um lado, membros do Círculo como Carnap e Neurath passaram a defender que enunciados são construções linguísticas que, como tais, só podem ser avaliados por parâmetros também linguísticos (Hempel 1935, 50). Enunciados e fatos são entidades distintas, e isso implicaria em que podemos avaliar a compatibilidade entre enunciados e analisar sua relação por meio de regras que produzam certas inferências, isto é, outros enunciados, a partir deles. Carnap passou a manter as ideias de Wittgenstein apenas a respeito da combinação entre enunciados para a formação de enunciados moleculares e da comparação entre eles sem se recorrer à realidade, já que os enunciados imediatos – seus enunciados "protocolares" – sobre a realidade deveriam ser aceitos imediatamente como verdadeiros (Hempel 1935, 51). Em suma, manteve-se a relação linguística entre enunciados, mas abandonou-se a estrita necessidade de correspondência com a realidade, também defendida no *Tractatus*.

De outra parte, e como consequência da aceitação dos limites que ligam a linguagem à realidade, os membros do Círculo reconheceram que o princípio verificacionista excluía, contraditoriamente à prática científica, todas as leis gerais do grupo de enunciados significativos, pois um enunciado de caráter geral não poderia ser verificado em sua totalidade. De acordo com a posição do *Tractatus* adotada pelo Círculo, um enunciado é verdadeiro quando retrata um estado de coisas existente na realidade, e dado o caráter universal e atemporal de uma lei empírica, entendida como uma afirmação molecular, seria impossível a verificação, em sua totalidade, de todos os enunciados atômicos que a compõe.

Com o convencionalismo adotado no final da fase verificacionista do Círculo de Viena, a linguagem da ciência poderia ainda manter os mesmos critérios lógicos para a determinação da validade de seus enunciados, mas a correspondência estrita com a realidade e a aceitação de uma única base empírica deixaria de representar o elo comum a todo o conhecimento. Afirmar a busca por um fundamento inabalável sobre o qual ancorar a teoria na realidade seria uma suposição muito forte, que estaria baseada na ideia de que existem propriedades absolutamente definidas às quais se poderiam referir inequivocamente. Carnap, por exemplo, reconhece que os próprios enunciados de experiência imediata devam ser revisados em determinados contextos em detrimento de outros enunciados que se mostrem mais bem fundamentados (Hempel 1935, 53). Desse modo, o valor de verdade

de uma proposição passaria a ser resultado não somente da correspondência com os fatos pertinentes a um determinado sistema teórico, mas dependeria, em algumas interpretações, quase que exclusivamente da coerência entre as proposições desse sistema.

Diante da abdicação progressiva da ideia de que a realidade é sempre acessada inequivocamente pela linguagem, a comunidade científica passar a desempenhar um papel fundamental no processo de validação dos enunciados empíricos. Isso leva à busca para se garantir, de maneira criteriosa, a confiança no conhecimento sobre a realidade. Para Carnap (Hempel 1935, 57-58), o desenvolvimento da ciência se deve ao acordo dos cientistas em torno da verdade de seus enunciados, à validação de sua adoção via experiência, e à consequente formulação de outros enunciados com base naqueles. A produção de enunciados verdadeiros é fruto de um treinamento que capacita os cientistas a elaborarem esses enunciados de uma maneira específica, mais precisa e adequada ao corpo da ciência. Ainda que diferentes sujeitos possam produzir diferentes enunciados, a comunidade que trabalha com a ciência deve se colocar em acordo sobre quais enunciados devem ser adotados em prol de outros enunciados e das teorias que se julguem mais relevantes.

Finalmente, o recorte que os empiristas lógicos fizeram do *Tractatus* no posicionamento inicial em torno do projeto de reestruturação da epistemologia impulsionou uma soma de esforços para renovação da filosofia e da lógica a partir da análise da linguagem. Uma das lições finais desse processo de "reconstrução" do conhecimento a respeito da ciência foi o de reconhecer que, ao mesmo tempo em que o mundo empírico não pode equivaler absolutamente a elementos linguísticos, enunciados também têm sentido para além do que se pode determinar empiricamente. Mesmo que Wittgenstein objetivasse outra direção para o *Tractatus* diferente daquela empreitada pelo Círculo de Viena, o contexto da época e a aproximação entre seus pensadores e suas idéias levaram a cabo, naquele momento, a tarefa que se incumbiram de realizar: a de buscar uma fundamentação do conhecimento.

Referências Bibliográficas

Barbosa Filho, B. (1981). Sobre o positivismo de Wittgenstein. *Manuscrito*, 5(1): 17-32.

Hahn, H., Neurath, O., Carnap, R. (1986). A Concepção Científica do Mundo "O Círculo de Viena". *Cadernos de História e Filosofia da Ciência*, vol.10, https://www.cle.unicamp.br/eprints/index.php/cadernos/article/download/1220/1011. Consulta: 26/05/2019.

Hempel, Carl. (1935). On the Logical Positivists' Theory of Truth. *Analysis*, 2(4), https://www.jstor.org/stable/3326781?seq=1#page_scan_tab_contents. Consulta: 26/05/2019.

Kenny, A. (1995). *Wittgenstein*. Madrid: Alianza Editorial.

Monk, Ray. (1995). *Wittgenstein: o dever do gênio*. São Paulo: Companhia das Letras.

Neurath, O. (2003). Fisicismo: A Filosofia no Círculo de Viena. *Philosophos* 8(1), https://revistas.ufg.br/philosophos/article/download/3209/3193/. Consulta: 26/05/2019.

Schlick, M. (1988a). O Fundamento do Conhecimento. In P. R. Mariconda (org.), *Coletânea de Textos Moritz Schlick, Rudolf Carnap*, pp. 65-81. São Paulo: Nova Cultural.

Schlick, M. (1988b). Sentido e verificação. In P. R. Mariconda (org.), *Coletânea de Textos Moritz Schlick, Rudolf Carnap*, p. 84. São Paulo: Nova Cultural.

Waismann, F. (1965). *The Principles of Linguistic Philosophy*. London: Macmillan.

Waismann, F. (1996). *Ludwig Wittgenstein und der Wiener Kreis: Gespräche*. Berlin: Suhrkamp Verlag.

Wittgenstein, L. (2001). *Tractatus Logico-Philosophicus*. São Paulo: Editora da Universidade de São Paulo.

Revista de Humanidades de Valparaíso
Año 7, 2019, 1er semestre, No 13, págs. 18-33
eISSN 0719-4242

Humanities Journal of Valparaiso
No 13 (2019): 18-33

On predicator rules and indexicality

Sobre las reglas de predicador e indexicalidad

Clément Lion

University of Lille, France
clement_lion@orange.fr

DOI: https://doi.org/10.22370/rhv2019iss13pp18-33

Recibido: 02/07/2019. Aceptado: 14/08/2019

Abstract

We argue that no attempt of reducing meaning to a systematic set of rules, according to which the role of linguistic expressions is to be normatively defined, can be abstracted from an irreducibly *decisional* compound. By comparing Lorenzen's project of building an *Ortho-language* (*Orthosprache*) and Brandom's *inferentialist* take on meaning, we distinguish two ways of acknowledging this fact, while claiming that Lorenzen's take is more genuinely constructive, insofar as *choices* be thought of as genuine features of constructions. It brings into a new perspective the relation between dialogical constructivism and Brouwer's intuitionism. Finally, we bring up a philosophical argument for the claim that interaction rules should be indexed on players and on their choices, when providing deontic bases to semantics.

Keywords: Ortho-Language, predicator rules, inferentialism, choices, dialogic, semantics.

Resumen

Argumentamos que ningún intento de reducir el significado a un conjunto sistemático de reglas, según el cual el papel de las expresiones lingüísticas debe definirse normativamente, puede abstraerse de un compuesto irreductiblemente decisional. Comparando el proyecto de Lorenzen de construir un lenguaje ortodoxo (*Orthosprache*) con el enfoque inferencialista de Brandom sobre el significado, distinguimos aquí dos formas de reconocer este hecho. Afirmaremos entonces que el enfoque de Lorenzen es más genuinamente constructivo, en la medida en que las elecciones se consideran características genuinas de las

construcciones. Esto nos llevará a una nueva perspectiva de la relación entre el constructivismo dialógico y el intuicionismo de Brouwer. De esta manera, presentamos un argumento filosófico para la afirmación de que las reglas de interacción deben indexarse en relación a los jugadores y sus elecciones, al proporcionar bases deónticas a la semántica.

Palabras clave: lenguaje ortodoxo, reglas de predicador, inferencialismo, elecciones, dialógica, semántica.

1. Introduction

It is nowadays widely accepted that the Augustinian picture of language —following Wittgenstein (1984)'s quotation of *Confessiones* I/8, which is intended at summing up the claim that words name objects— cannot provide a satisfying account of the way meaning is constituted. One alternate proposal to such a picture consists in thinking meaning as being *immanent* to the (correct) use of language, following Wittgenstein's alleged slogan, namely that "meaning is use". *Inferentialism* represents currently a research program whose aim is to give a semantic account of any part of language in terms of the way it contributes to carrying out inferences, to which a speaker (tacitly) commits himself when uttering a sentence. This research program stems partly from some of Sellars' insights — in particular that "to grasp a concept is mastering its use", but it would be fair to say that, despite of some differences, it relies also to perspectives that were deployed within the frame of the so called *Erlangen Constructivism*. In Kamlah and Lorenzen (1996, 73), the meaning of a predicate is explicitly defined in terms of determinate relations that are associated to a set of tacit commitments a speaker is to be responsive for when speaking. Understanding the meaning of a predicate is nothing else than knowing which commitments one takes when uttering something and according to which rules these can be assessed. It corresponds thus closely to the "game of giving and asking for reason" (Brandom 2000, 165) whose practical mastery represents, Brandom says, what understanding the meaning of an utterance actually consists in, even though, as we will see, there might be crucial shades between the respective concepts of a *predicator rule*, in Lorenzen's terminology, and of an *inferential relation*.[1]

Now, that meaning be immanent to the use of language, in such a way that *referring to* be explained in terms of *inferring* does not preclude, of course, that the use of language be itself indexed on something extra-linguistic: the non-linguistic circumstances of statements are inferentially articulated to them, in so far as, by being confronted with certain perceptual contents, a speaker is committed to make (or, at least, not to make) determinate assertions. The source of the inferential normativity that relies non-linguistic circumstances to utterances is the self-constitution of a linguistic community as setting up

[1]According to Rahman et al. (2018, 268-271), the former should rather be conceived as a *transition relation*: it expresses a *presupposition*, which is not the same as a genuine inferential relation.

the norms of meaningfulness (Brandom 1994, 119). In Brandom's view, adopting a (dynamic) system of rules depends thus on integrating a community for which these rules are holding good: it is thus a matter of going along the lines of a common institution. Hence, meaning relies to a *normative fact*, namely the fact that "one uses words this way". It does not mean that explicit rules can ever be stated once and for all, as representing the correct way of using words, but it means that any linguistic community regulates itself in such a way that the correctness of uses is dynamically laid down by the community itself and that being submitted to these norms of correctness presupposes entering into a community as a responsible speaker.

Now, let us observe that were there several different linguistic communities, submitted to different systems of rules, however there should be also a common principle of regulation. Ultimately, no local space of meaning —even taking the form of a whole historical community— may ever be considered as representing something else as a provisional degree of explicitness. In other words, the normative fact that consists in contingently belonging to a determined community should tend to be diluted in a normative *universal* realm, located beyond any factual anchoring of a concrete speaker. In Brandom's words:

> Each one defines a different way of saying 'we', each kind of 'we'-saying defines a different community, and we find ourselves in many communities. This thought suggests that we think of ourselves in broadest terms as the ones who say 'we'. It points to the one great Community comprising members of all particular communities - the Community of those who say 'we' with and to someone, whether the members of those different particular communities recognize each other or not. (Brandom 1994, 4)

2. Inferentialism and subjectivity

If one assumes indeed an inferentialist perspective on meaning, and especially on the meaning of elementary propositions, then one must also assume that any use of a predicate should be normatively regulated, at least through the implicit preclusion of all possible incorrect moves involving it, out of a common (and ideally universal) space of use. Now, within the space which is thus negatively defined, an innumerable quantity of moves are still permitted. Each of those participates to specify the stitches of a common web of meaningful statements,that represent potential discursive positions, insofar as they enable one to keep the line against specification requests, possibly addressed to any utterer taking part of it. This is certainly the way scientific language is regulating its own use, by trials and errors.

Someone who would like to go on speaking out of this regulative process will be soon disqualified as talking non-sense, or at least as talking alone: he will locate himself out of the common linguistic space as it is normatively instituted. It does not mean at all that no singular creation of meaning be ever thinkable: it only means that once created, then

meaning has to be regulated through collective use, so that only these statements, that are conservative over other already well accepted statements in their most far implications and presuppositions —unless they indicate explicit ways to transform some of these well accepted positions into more efficient ones—, will be able to keep on going in the public space of meaning. Accordingly, any meaning assumption is to be thought of as being essentially communicable and potentially strengthened through the process of communication.

Therefore, any strictly *subjective* position should be thought of as tending to dilution in a public space of assessment. In other words, the meaning of concepts is to be thought of as a (universally) normative feature (Brandom 2000, 29). For several potential utterers, understanding one and the same concept in different ways would be thus only a matter of being more or less close to the grasp of the whole web of inferential relations within which it is to be located. There are several degrees of mastery: no conflicting ways of *rightly* grasping a concept may be sustained *on the long run*, unless its meaning is still on the way to be determinate. One may factually grasp only a small part of the relevant inferential relations, according to the anchoring of his life within a particular context of use, but the very norm according to which his own speech is to be assessed is ultimately beyond this factual grasp. Accordingly, *one may understand incorrectly the meaning of his own utterances*: subjectivity is not the right location of meaning and no one is entitled to claim that his own understanding be more accurate than the public assessment of it, at least not without a rightly defensible justification, eventually detachable from him.

If the meaning of utterances depends on mastering their correct use, then assessing this correctness presupposes making the rules explicit, according to which these utterances play a role within concrete practices. Now, the process of making them explicit necessarily depends on *choosing* a device through which projecting concrete uses on a (formalized) means of expression of the accorded rules becomes possible, in such a way that the correctness of an inferential relation can be communally assessed, by referring to determinate criteria one has agreed on. Hence, meaning has to be *artificially* identified with an ideal set of norms enabling one to give an account of every compounds of meaningful utterances.

In this extent, even though inferentialism is explicitly connected with Wittgenstein's second philosophy, it would be misleading to reduce the latter to the former, because Wittgenstein explicitly rejects any claim of possessing such a unified device to approach meaning (Wittgenstein 1984, 250). Beyond his admittance of *given* forms of life on which every possible uses of signs are to be anchored, a further (decisional) step has to be carried out in order to build a *systematic* approach such as inferentialism aims to be. Indeed, it presupposes admitting a certain way of reconstruction as reliable for clarifying *any* meaning. Namely, inferentialism is based on linking meaning to argumentative justification, and therefore it is based on promoting a particular way of using utterances as a

key concept for understanding every other ways. This particular way namely consists in projecting concrete speeches into the aforementioned "game of giving and asking for reasons", that is to say into something they seem not always to be at first sight.

According to Brandom, such a *particular* reconstruction of the way meaning *could be* constituted is not to be understood as the only legitimate one. His aim is rather to show how it enhances the ability to explain the very nature of meaning (in contradistinction with Wittgenstein's "theoretical quietism", following Dummett 1978a's expression).[2] The (theoretical) decision on which inferentialism is based is such that admitting its legitimacy depends on taking it over, i.e. on projecting oneself into the space which comes to be instituted by it. In other words, a factual choice comes before theorizing and then *dilutes itself* in the objectivist normative view on meaning that it has been instituting.

As a result, no subjective norm of meaning may ever be sustained, even though subjective (but as such "misleading") uses make the most part of real speeches. Indeed, new conceptuality may arise from a particular stance, but it would become graspable only if strategical inferential paths are cleared from the chaos of the awkward attempts through which concrete speeches mostly collapse. The institution of meaning through common regulation depends on locating any claim to meaning within a common game of giving and asking for reason. This game is not thought of as being anchored in a particular form of life, but rather in the very fact of the (provisional) plurality of forms of life, as naturally tending to communality and unity. One's particular refusal to submit one's own claim to meaning to a common regulation device may represent a particular stance, but it could never rely to a fully understandable meaning articulation, because it would always hit boundaries depending on arbitrary choices that another speaker might not be willing to take over. In this view, meaning should always be communicable and communicability should be itself reduced to argumentative defensibility.

But such an equation may not be completely necessary; the choice of a normative system might be preferably seen as not being *repressed* out of the frame of any argumentative stance, as we will see in the next section, but rather *expressed* as a constructive compound of it.[3] The question is whether a subjectively defined frame of speech has necessarily to be doomed to *semantic insignificance* when it refuses its own dilution in public assessment?

---

[2]"The idea is to show what kind of understanding and explanatory power one gets from talking this way [i.e., from the inferentialist sight], rather than to argue that one is somehow rationally obliged to talk this way." (Brandom 1994, xii). Now one may wonder whether it really makes a difference: if one convincingly shows that the adoption of this sight actually increases the explanatory power, than refusing it will represent an irresponsible stance.

[3]Brandom acknowledges a *decisional* compound in adopting the inferentialist framework, but once it is adopted, then one is committed to saying that any meaning is ultimately to be thought of as depending on the way it contributes to winning in the game of giving and asking for reason. Any stance that makes a speaker fail in this game must ultimately fade out as a claim to meaningfulness. In this extent, meaning only counts from a strategical sight, as we will see below.

3. The program of building an Ortho-language (*Orthosprache*); differences with the program of Inferentialism

As it has been initially formulated Kamlah and Lorenzen (1996) and Lorenzen and Schwemmer (1975), the program of building an *Orthosprache* (or "Ortho-Language") consists in reconstructing, step by step and from the scratch, the language of science by anchoring any meaningful sort of expression in a context of use, whose description requires only very elementary forms of articulations. It leads also to distinguishing sorts of words, so that anyone is enabled to make each part of an utterance associated to a teachable operation such that any complex statement holds as the result of a constructive path through which an elementary operation is developed step by step into a complexer one. Thus, no gap remains that would hinder a full mutual understanding. A *predicator rule* represents a means to make fully explicit the operative bases on which a predicator is to be defined, starting from a broader one and specifying its own way of operating.

In despite of similarities, Brandom's inferentialist take on meaning and Lorenzen's one diverge on a crucial point. While the former represents a theoretical option, aiming at giving an account of what meaning generally *consists in* (while conceding that, eventually, other explaining options are possible), the latter is based on acknowledging a factual discrepancy between many ways discourses are articulated as meaningful, so that building an Ortholanguage represents a task to be done, if one wants to unify scientific community. In other words, this Ortho- language must not be thought of as a theoretical means to make explicit what remains implicit when people scope at mutual understanding, but rather as a practical undertaking, intended to create the concrete conditions for building a unified community of speakers: usual ways of speaking *are not* essentially submitted to predicator rules! In Brandom's view, a universal norm of meaning is supplied through the device of argumentative praxis; in Lorenzen's view, such an universal norm is not *given*, but should be *artificially created*. It thus means that the way meaning is actually constituted is much more dependent on factual contexts of use in Lorenzen's view than in Brandom's. Even the very project of building an Ortho-Language depends on an historical context in which the discrepancy of uses sterilizes most of the philosophical investigations, insofar as it makes every theoretical product relative to a particular school, isolated from every others (Kamlah and Lorenzen 1996, 1).[4]

The *Ortho-Language* that Lorenzen was aiming at, as it is to be especially deployed through predicator rules in association with situations of teaching and learning, must be considered as an *artificial device*[5]*:* local discourses, essentially based on particular

[4]In the *Logische Propädeutik*, the concepts that express themselves in the natural language are not submitted to *predicator rules*, but rather to contextual uses, such that *plasticity* is one of their most essential features. It does not preclude however their own meaningfulness, inasmuch as they contribute to the opening of a world. In natural language, there is no necessity to go beyond this opening in order to deploy meaningful stances through language. *Language* must not necessarily be *Ortho-Language*.

[5]It is not the same as *artificially* identifying meaning to inferential relations: in one case, we have a *theo-*

empragmatical anchorages, still belong to the domain of sense, in Lorenzen's view. By contrast, from an inferentialist sight, the set of the inferential relations according to which any sentence deploys its genuine meaning, must be conceived as an ideal norm of correctness, not depending on any decision taken by concrete speakers. At most, it is the adoption of the inferentialist framework, as a good theoretical device, challenging other possible ones, which may be eventually reported to a decision, but in such a way that it should ultimately be argued for in the game of giving and asking for reasons. In case this device reveals itself as the better one *on the long run*, there is no reason to accept that any meaning could be constituted out of such a common space of assessment.

4. On two ways of rejecting incommunicability

The very idea of building an *Orthosprache* is deeply connected with the rejection of any call to privacy in the constitution of a rational discourse, especially when dealing with sciences and mathematics. However, it seems not to be equivalent to saying that, in Lorenzen's view *meaning itself* should preclude privacy. It is rather the case that every day speeches are anchored in concrete situations, depending on inexpressible factual parameters, that are irreducible to any abstract device on which one could base to make them universally accessible. Any one finds himself within a determinate language whose meaning cannot be, as logicians tend to claim, abstracted from its historical situation, without falling into vagueness. In fact, when located in this situational anchoring, "vagueness" holds as a positive feature, that Kamlah and Lorenzen rather call "plasticity" (Kamlah and Lorenzen 1996, 65).

This is an important point to be made if one considers the way Lorenzen speaks of Brouwer's challenge to classical logic. In *Logik und Agon*, he describes it in term of a situation such that, suddenly, a group of people sees one of the most deeply anchored principles of logic, namely the excluded-middle, as unreliable (Lorenzen and Lorenz 1978, 2). This sudden arising of a mathematical practice finds its own distinctive expression in refusing to endorse the kind of (formal) inferential commitments, that have been traditionally formalized by logicians. It breaks accordingly the bridge between these "deviant" mathematicians and the community of those who are sticking to classical forms of reasoning. Even though this new mathematical practice have, at least provisionally, less demonstrative power than the usual one, one cannot assume that the utterances that are anchored in it must be denounced as meaningless. It is rather the case that a new stance has arisen through intuitionism, associated to the claim of its own irreducibility to the space of assessment of mathematical truth which were holding until then and which is based on the idea of a truth in itself. Such a break off was carried out through Brouwer's singular and unconventional voice. It has been given, to that extent, a philosophical basis

retical artifice, in the other case an *institutional* one, insofar as it puts a practical program on the table, that could easily be transformed into a political one.

by Becker (1927, 636), to whom intuitionism represents a decisive step towards the acknowledgement of "facticity" within mathematics, contrarily to formalism, that ignores it. Becker's sight is indeed based on an Heideggerian conceptuality, yet such that the concerns with "authenticity" are now related, as it is *not* the case by Heidegger, to mathematical activity. Accordingly, one of the most controversial devices, namely the "free choice sequences", by which Brouwer attempted to give an account of what continuity is, expresses now, in Becker's mind, one of the most authentic mathematical construct (even though it tends, precisely because of its authenticity, to contradicts its own *mathematical/formal* essence) (Becker 1927, 674). Accordingly logic, even though defined as a system of rules no justification should be asked for — because it allegedly delivers the basic material that anyone needs in order to supply any sort of justification — cannot hold any more as an impassable universal medium, out of which no meaningful discourse can ever be articulated. Formalist discourses, though consistent, miss their own factual anchoring in decisional parameters — dealing with time and historicity in Becker's perspective — that they tend to forget.

In Lorenzen's view, logic must not be seen as a universal normative frame prefiguring every possible meaningful utterance. In league with Becker, he rather sees it as a device to which a function must be assigned in relation to the aim it is intended to serve. For both of them, purely formal logic is a way to flew away from facticity towards a normative stance, that enables one to push away its own (constructive) responsibility. However, Lorenzen clearly diverges from Becker when formulating the program of building an *Ortho-Language*. Even though *meaning* is not to be regulated through an allegedly universal abstract system of more or less broaden rules, the dialogue between different meaningful positions should regulates itself in such a way that a collaborative (re-)construction of a common system of rules holds, for any local speaker, as a possible factual choice, made from within his own local realm. Accordingly, the rejection of incommunicable features out of the domain of sense holds here as a *practical* commitment towards a dialogically regulated common frame whose construction depends on each speaker taking over the same type of commitment, in order to carry out a common basis for future constructions, but also on each speaker showing what eventually prevents him from adhering to this basis.

This way is to be carefully distinguished from the one that stands at the basis of Brandom's inferentialism and that was already expressed by Dummett:

> To suppose that there is an ingredient of meaning which transcends the use that is made of that which carries the meaning is to suppose that someone might have learned all that is directly relevant when the language of a mathematical theory is taught to him, and might then behave in every way like someone who understood that language, and yet not actually understand it, or understand it only incorrectly. (Dummett 1978b, 217-218)

It is the very same idea that is expressed by Brandom, in a broadest sense, when saying that

> […] only what plays a suitable role in essentially social, indeed linguistic, discursive deontic scorekeeping practices should count as conceptually contentful in the fundamental sense. (1994, 608)

That purely private compounds have to be rejected out of the domain of meaning is not to be understood, as it is the case by Lorenzen, in terms of a *constructive decision*, but rather in terms of a *semantic impossibility*, related to an objectivist view on what semantic is to be. It is ultimately based on the idea that no discursive commitment may ever be taken by a speaker without ascribing to other speakers the ability to assess whether it will have been correctly fulfilled (which depends on knowing unambiguously which inferential relations apply). The existence of a common normative frame of assessment is accordingly a necessary presupposition of any claim to meaning: this existence does not depend on any decision; it plays the role of a *transcendental* condition of meaning.

In Lorenzen's view, no such presupposition should be made without indicating how a constructive path towards it will be defined. Therefore, the decision of rejecting incommunicability out of the space of meaningful utterances cannot produce anything more than an essentially provisional common stance, which will always be threatened by the possible arising of new diverging practices disturbing mutual understanding in its most uncontroversial anchoring. If the *Ortho-Language* aims at rendering any articulation of scientific language in a dialogically constructive way, then it must give a place to meaning *creation*, as it may consist in destabilizing any fixed constructive articulation without necessarily destabilizing the accorded syntax. Accordingly, there must be a place for sense shifting, even when sense is decidedly located on a common constructive basis.[6]

[6]This is a point that is been investigated into by Schneider (1999) and Schneider (2014). While arguing in league with Wittgenstein that the program of unifying syntax and semantics within a systematic *theory of meaning*, defined in terms of rules that must regulate any type of linguistic use, is doomed to failure, Schneider refuses to endorse a "dismissive quietism", which would consist in sterilizing from the scratch any program of systematizing meaning. His claim is that language has a "systematic side", making its use similar to calculating by following rules, but it is such that it gets its own meaningfulness through *another side*, requiring "imagination". While no formalization will ever enables one to get through all possible meanings an utterance may have, because "grammatical forms are continually *projected* into new fields of activities" (Schneider 2014, 170) it represents however a key-device, without which no mastering of a language will ever be possible to acquire. Additionally, Schneider (1999, 472-493) argues that the program of building an Ortho-Language by anchoring every compounds of meaning in a typical situation, which makes it unambiguously teachable, cannot avoid a diverging of syntax and semantic to arise, due to the essential variability of the ways of acting that may be embedded into one and the same syntactic form of presentation.

5. Can a privately deviant use of a given predicate be communicable?

In order to explain what we have here in mind, we may take an example. Let us assume that an expert in a domain, for instance a skateboarder, tries to teach how to make a particular trick to a beginner. While saying "you must pop your deck this way", he shows exactly what he means and repeats the move several times until the beginner identifies precisely the relevant compounds of the move. Let us assume that the latter is now able to identify when himself, or anyone else, is rightly doing the move. It seems obvious that there will certainly remain ingredients, transcending the inferential uses of the expression "I pop my deck", and however essential to the correct understanding of it. It is certainly the case, insofar as the trick at stake takes place in a determinate *form of life*, that consists in "being a skateboarder".

The same may certainly be said concerning mathematics, as it has been deeply observed by Le Roy (1930, 118-121). Deploying a Bergsonian stance on creativity within mathematical practice, Le Roy argues that without a "driving image" ("image motrice"), no definitional rule can ever succeed in realizing a content for operative thinking. No constructive path can ever be followed if no pre-discursive image drives it, while furnishing more than what can ever be made explicit, namely the principle of a continuous constructive move, which also essentially *cannot* be made explicit, because analysing it is tantamount decomposing its continuity into separate moves. Mathematical educators should, Le Roy says, accordingly feed their pupils with something that is usually totally neglected, namely the "imaginative vitality". It is now certainly the case that such a Bergsonian conceptuality is burdened by a *representationalist* stance on meaning, as broadly denounced by Brandom (2000, 45-47): it presupposes an immediate access to an ineffable mental content, i.e. a private ingredient mysteriously enshrouding the use that an individual makes of mathematical symbols. It would certainly fall under the aforementioned Dummett's attack on allegedly private compounds of meaning. Our aim here is certainly not to contest the well-foundedness of such an attack. There is no doubt that it makes no sense to assume an absolutely private ingredient which would be nonetheless essential to mathematical understanding. The point is not yet exactly such, because Le Roy puts specifically the emphasis on the problem of *transmitting* this imaginative vitality: it presupposes thus an acknowledgement of its ability to take place within a common space of activity, eventually according to (at least informal) *rules of interaction*.

Brouwer was certainly deeply concerned with such a kind of matters, for instance when reporting his meeting of G. Mannoury:

> As happens so often, I began my academic studies as it were, with a leap in the dark. After two or three years, however full of admiration for my teachers, I still could see the figure of the mathematician only as a servant of natural science or as a collector of truths: — truth fascinating by their immovability, but horrifying by their lifelessness like stones from barren mountains of disconsolate infinity. And as far as I could see there was room in the mathematical field for talent and devotion,

> but not for vocation and inspiration. Filled with impatient desire for insight into the essence of the branch of work of my choice, and wanting to decide whether to stay or go, I began to attend the meetings of the Amsterdam Mathematical Society. There I saw a man apparently not much older than myself, who after lectures of the most diverse character debated with unselfconscious mastery and well-nigh playful repartee, sometimes elucidating the subject concerned in such a special way of his own that straight away I was captivated. I had the sensation that, for his mathematical thinking, this man had access to sources still concealed to me or had a deeper consciousness of the significance of mathematical thought than the majority of mathematicians. [His papers] had the same easy and sparkling style which was characteristic of his speech and, when I had succeeded, not without difficulty, in understanding them, an unknown mood of joyful satisfaction possessed me, gradually passing into the realization that mathematics had acquired a new character to me. For the undertone of Mannoury's argument had not whispered: "Behold, some new acquisitions for our museum of immovable truths", but something like this: "Look what I have built for you out of the structural elements of our thinking. — These are the harmonies I desire to realize. Surely they merit that desire? – This is the scheme of construction which guided me. — Behold the harmonies, neither desired nor surmised, which after the completion surprised and delighted me. – Behold the visions which the completed edifice suggests to us, whose realization may perhaps be attained by you or me one day." (Brouwer 1975, 474-475)

One is confronted here with private ingredients of meaning —or better said "of significance" whose privacy should not be considered, however, as an absolute feature. They represent, so to say, a *communicable privacy*, whose communicability depends on parameters that do not let themselves easily embed into a well-defined situation of teaching and learning, because they depend on blurred kinaesthetic wholes, whose communality can very hardly be willingly constructed.

Let us call this type of use "privately deviant", because the private compound at stake is based on very small deviations from observable rightness, such that it lets itself hardly reach by discursive reports. These *semantic* deviations might be, in fact, compared with the very structure of the continuum, as thought of within Brouwer's mathematical intuitionism (Brouwer 1928, 4). If one takes into account the distinctive feature of the free choice sequences, namely their essential openness to further determination through a potentially infinite process of choice, then their being different is not equivalent to being apart one from another. Similarly, a particular use of mathematical language might be enshrouded by a "whispering undertone", making no explicitly reportable difference, while opening a space of play within which meaning, as regulated by use, would let itself *index* on unforeseeable differences, that reveal themselves in response to accurate requests, prescribed by no explicit rule, but also not restricted by the rules at work. By referring to a such similarity between a privately deviant compound of meaning and the open struc-

ture of a free choice sequence, we may call *semantic continuum* the domain of analysis that an extended theory of meaning should investigate into, when aiming at rendering the dynamic features of meaning, that look like Mannoury's "whispering undertone".

The question is now: how to embed *infintely proceeding sequences* within dialogues in order to structure such a semantic domain of analysis, in such a way that a bridge can be built between *meaning*, as understood in this extended way, and *truth*, which any mathematical practice, even in Brouwer's mind, will always be aiming at?

6. On the *play level* as a key device to make sense of what stands beyond the reach of Inferentialism

We shall present here, in a purely programmatic way, a possible use of the crucial distinction, initially pointed out by Kuno Lorenz, within the framework of *dialogical constructivism*, between the play —and the strategic— levels (Lorenz 2010). It is important to point out that it represents certainly a continuation of Lorenzen's own sights, even though the emphasis is now differently distributed.

Let us recall that the kernel of the dialogical view on semantics consists in the idea that the meaning of an utterance is given by the way a justification process can be supplied through the execution of a *finite* dialogue starting from it, according to determinate *interaction rules*, acknowledged as such by a Proponent (P), who defends it, and an Opponent (O), who challenges it. No basis for justification has to be looked for outside the dialogue itself. No matter which *private* content one locates behind an assertion, the only relevant point is that by making this assertion, he entitles his interlocutor to tackle it according to the rules (which are at least of two sorts: *particle rules*, which are *player-dependent* and *structural rules*, which takes the nature of the player into account).

Now, the play —and the strategic— levels correspond to different (but complementary) sights on dialogues. From the first perspective, dialogues are executional processes, depending on factual choices that are made by concrete players. To that extent, the fact that P wins a dialogue cannot hold as a proof that his claim is true, because it may depends on O's (bad) choices. In order to make sense of what truth is, on must step into the second perspective, which depends not any more on the players' factual choices. In Lorenz's words:

> […] it is useful to observe that win and loss of a dialogue about a given proposition will in general depend upon an individual play of the game and will not be a function of the proposition alone. But the strategies of either player of the game are invariant against the choice of arguments of the other player. Hence, a proposition A shall be called 'true', iff there is a winning strategy for A; this means that the player who is asserting A —the proponent P— will be able to win a dialogue on A independently of the choice of arguments of the opponent O. (Lorenz 2010, 11)

We shall observe now that Brandom's inferentialist take on meaning corresponds to a *strategic reading* of dialogues, because it is based on stating the very existence of inferential relations, whose normative feature is not thought of as depending on the choices that are factually made by concrete interlocutors. But, as Rahman et al. (2018, 270) clearly points out:

> *not every sequence of moves in games of asking for reasons and providing them is necessarily inferential, only those plays leading to winning strategies are.*

By sticking to the (universally) normative compound of meaning, Brandom's inferentialism is not any more than Dummett in position to extend the domain of semantic by incorporating into it, the aforementioned ingredient that transcends the normatively regulated use of any utterance, namely the one we called *privately deviant*. But now, if we go down into the play-level, at which factual choices are relevant, nothing prevents us any more from specifying rules that will regulate the factual use of ingredients, that may reveal themselves as totally irrelevant from a purely strategic point of view, while being however *significant*. These rules may take into account indexical compounds, if we are to take over Peirce's manner of calling the structure of a sign that relies to its object because of its being really affected by it (Peirce 1974, II. 248). Such ingredients as kinaesthetic wholes could accordingly take, as indexes or indexical processes, part of a common space of play, regulated by rules, not to be confused with objective norms in front of which any strictly subjective stance has ultimately to dilute itself. These indexical processes might be mathematically structured by using Brouwerian choice sequences as a form of presentation.

In other words, we suggest exploring, along the lines of Lorenz's crucial emphasis on the *play level*, what we called the *semantic continuum*, by *mathematically* approaching the rules through which singularities seek mutual understanding, even when they are assigned to formally determinate roles within strictly normative forms of communities. To that extent, we are here following a Brouwerian path, which is however clearly not solipsist, because it is constitutively dialogical. Indeed, the aim of building an Ortho-Language might incorporate its own immanent creative and metaphorical processing.

7. Conclusion: a plea for indexing rules of inter- action on players and choices

In his recent Oslo and Stockholm lectures, Per Martin-Löf has also taken a dialogical route towards a foundation of logic, and language in general, on *ethics* or *deontology* (Martin-Löf 2017a; Martin-Löf 2017b). In this perspective, the rules that regulate the use in which meaning consists should be called *rules of interaction*, rather than *rules of inference*. That a speaker be committed to assert *C'* because he has asserted *C*, depends ultimately of his being requested to by a hearer. Thus, in its most general form, the rule, which the meaning of an assertion is to be based on, could be schematically rendered in the following way:

$$\frac{\vdash C \qquad ? \vdash_{may} C}{\vdash_{must} C'}$$

It indicates that, while asserting *C* a speaker implicitly gives to any hearer the right of *requesting* a justification for it, which represents accordingly a duty for himself.

Now, as observed by (Rahman et al. 2018, 293), it is not exactly the way a dialogician would render it. It would rather take the following form:

$$\begin{array}{c} \vdash^{X} C \qquad ? \vdash^{Y}_{may} C \\ \Downarrow \\ \vdash^{X}_{must} C' \end{array}$$

Firstly, instead of a bar, we should use an arrow, to remind Lorenzen's action-based background which is to be carefully distinguished from a purely inferential rule, as essentially disconnected from its operative anchoring. It is not only a question of notation: the bar relies to a norm that prescribes a duty, while the arrow rather indicates the possibility of interacting this way, which is accordingly a matter of choice. One can identify and apply such a scheme of interaction in the same way as one may learn how to knit or build a wall: in every case, the schemes at stake represent possible constructive path, and not categorical imperatives. Such a preference of dialogicians in choosing the arrow sign may find a justification in what we said before. The institution of a commonly regulated space of play essentially depends on the choice we make of constructing it and incorporating it into our forms of life. As such, a meaningful stance might aim at keeping out of this collective program.

It directly relies to what Rahman *secondly* points out, namely that this interaction rule would be, from the usual perspective of a dialogician, indexed on players (*X* and *Y*). We go now a little further on, by arguing that it *must* be indexed on players. Our argument relies on what we said previously, namely that no justification for constructing a common space of play can be given without observing fundamental divergences in singular ways of using language and of anchoring it in concrete situations. Martin-Löf's manner of symbolizing this fundamental rule of interaction misses the crucial point by staying disembodied from the very fact that forms of life are plural. In other words, he sticks, as a logician, to a formal view of what a social interaction is, namely that it is functionally definable and that the singular series of choices that institute each speaker as such brings nothing essential to the way functional relations are to be defined. By starting from a more radical point of view, we claim that these singular series of choices constitute the reason why something like *propositional functionality* in general emerges, namely as a constructive means to build a common space of play *against* the irreducible *fact* of incommunicability.

References

Becker, O. (1927). *Mathematische Existenz*. Tübingen: Niemeyer.

Brandom, R. B. (1994). *Making it explicit. Reasoning, Representing and Discursive Commitments*. Cambridge, Massachussets and London: Harvard University Press.

Brandom, R. B. (2000). *Articulating Reasons. An introduction to Inferentialism*. Cambridge, Massachussets and London: Harvard University Press.

Brouwer, L.E.J (1975). *Collected Works*. Ed. by Arend Heyting. Amsterdam: North-Holland.

Brouwer, L.E.J. (1928). *Die Struktur des Kontinuum*. Wien: Komitee zur Veranstaltung von Gastvorträgen ausländischer Gelehrter der exakten Wissenschaften.

Dummett, M. (1978a). Can Analytical Philosophy be Systematic, and ought it to be? In *Truth and other Enigmas*. Cambridge: Harvard University Press.

Dummett, M. (1978b). The Philosophical Basis of Intuitionistic Logic. In *Truth and other Enigmas*. Cambridge: Harvard University Press.

Kamlah, W., Lorenzen, P. (1996). *Logische Propädeutik: Vorschule des vernünftigen Redens*. Stuttgart, Weimar: Verlag J. B. Metzler.

Le Roy, E. (1930). *La Pensée intuitive, II. Invention et Vérification*. Paris: Boivin et Cie.

Lorenz, K. (2010). *Logic, Language and Method*. Berlin: De Gruyter.

Lorenzen, P., Lorenz, K. (1978). *Dialogische Logik*. Darmstadt, Germany: Wissenschaftliche Buchgesellschaft.

Lorenzen, P., Schwemmer, O. (1975). *Konstruktive Logik, Ethik und Wissenschaftstheorie*. Mannheim: Bibliographisches Institut A.G.

Martin-Löf, P. (2017a). Assertion and request. Lecture held at Oslo. Transcription by A. Klev. (no published work)

Martin-Löf, P. (2017b). Assertion and request. *Lecture held at Stockholm*. Transcription by A. Klev. (no published work)

Peirce, Ch. S. (1974). *Collected papers of Charles Sanders Peirce*. Ed. By Charles Hartshorne and Paul Weiss. Cambridge, Massachussets: The Belknap Press of Harvard University Press.

Rahman, Sh. et al. (2018). *Immanent Reasoning or Equality in Action: A Plaidoyer for the Play Level(Logic, Argumentation & Reasoning Series)*. Dordrecht: Springer International Publishing.

Schneider, H. J. (1999). *Phantäsie und Kalkül. Über die Polarität von Handlung und Struktur in der Sprache*. Frankfurt am Main: Suhrkamp.

Schneider, H. J. (2014). *Wittgenstein's later Theory of Meaning*. Trans. by Timothy Doyle and Daniel Smyth. Chichester, UK: Wiley Blackwell.

Wittgenstein, L. (1984). *Philosophische Untersuchungen*. Frankfurt am Main: Suhrkamp.

Revista de Humanidades de Valparaíso
Año 7, 2019, 1er semestre, No 13, págs. 34-67
eISSN 0719-4242

Humanities Journal of Valparaiso
No 13 (2019): 34-67

Hybrid Logic as Extension of Modal and Temporal Logics

La lógica híbrida como extensión de las lógicas modal y temporal

Daniel Álvarez-Domínguez
Universidad de Salamanca, España
danielalvarezdguez@gmail.com

DOI: https://doi.org/10.22370/rhv2019iss13pp34-67
Recibido: 16/07/2019. Aceptado: 14/08/2019

Abstract

Developed by Arthur Prior, Temporal Logic allows to represent temporal information on a logical system using modal (temporal) operators such as P, F, H or G, whose intuitive meaning is "it was sometime in the *P*ast...", "it will be sometime in the *F*uture...", "it *H*as always been in the past..." and "it will always *G*oing to be in the future..." respectively. Valuation of formulae built from these operators are carried out on Kripke semantics, so Modal Logic and Temporal Logic are consequently related. In fact, Temporal Logic is an extension of Modal one. Even when both logics mechanisms are able to formalize modal-temporal information with some accuracy, they suffer from a lack of expressiveness which Hybrid Logic can solve. Indeed, one of the problems of Modal Logic consists in its incapacity of *naming* specific points inside a model. As Temporal Logic is based on it, it cannot make such a thing neither. But First-Order Logic does can by means of constants and equality relation. Hybrid Logic, which results from combining Modal Logic and First-Order Logic, may solve this shortcoming. The main aim of this paper is to explain how Hybrid Logic emanates from Modal and Temporal ones in order to show what it adds to both logics with regard to information representation, why it is more expressive than them and what relation it maintains with the First-Order Correspondence Language.

Keywords: Arthur Prior, First-Order Correspondence Language, Temporal Representation, Nominals, Translations.

Resumen

La lógica temporal fue creada por Arthur Prior para representar información temporal en un sistema lógico mediante operadores modales-temporales como P, F, H o G. Intuitivamente tales operadores pueden entenderse respectivamente como "fue alguna vez en el pasado...", "será alguna vez en el futuro...", "ha sido siempre en el pasado..." y "será siempre en el futuro...". La evaluación de las fórmulas construidas a partir de ellos se lleva a cabo en semánticas kripkeanas y, de este modo, la lógica modal y la temporal están relacionadas. Sin embargo, aunque sus mecanismos permiten formalizar la información modal-temporal con cierta precisión, ambas lógicas adolecen de un problema de expresividad que la lógica híbrida es capaz de solventar. En efecto, uno de los problemas de la lógica modal reside en su incapacidad para *nombrar* puntos concretos dentro de un modelo. La lógica temporal, al basarse en ella, tampoco puede hacerlo. Pero la lógica de primer orden sí es capaz gracias a las constantes y a la relación de identidad. La lógica híbrida, que resulta de combinar la lógica modal con la lógica de primer orden, sería una solución a este problema. El principal objetivo de este artículo consiste en explicar el origen de la lógica híbrida a partir de la modal-temporal para mostrar qué añade a ambos sistemas en la representación de información, porqué es más expresiva que ellos y qué relación guarda con el lenguaje de correspondencia de la lógica de primer orden.

Palabras clave: Arthur Prior, lenguaje de correspondencia de primer orden, representación temporal, nominales, traducciones.

1. Introduction

Arthur Prior (1957) (1967) (2010) developed one of the most classical and fundamental approaches of Temporal Logic (TL), which can be understood in general terms as the representation of propositions with temporal information on logical frames.

In Classical and Modal Logic we are used to deal with formulae of the kind ¬p, p → q, □p or ◇p, among others, where p and q are variables which stand for whatever propositions, → is the material conditional and □ and ◇ are modal operators standing for necessity and possibility respectively.

Modal Logic syntax is not much wider than Classical one. We just have to add the rest of connectives. Its semantics, on the other hand, is based on the notion of *possible worlds* to valuate formulae such as □φ or ◇φ: □φ is satisfied at a world w of a model $\mathfrak{M}$ if and only if is satisfied at every world w′ accessible in $\mathfrak{M}$ from w, whereas ◇φ is satisfied at a world w of a model $\mathfrak{M}$ if and only if there exists a world w′ in $\mathfrak{M}$ accessible from w where φ is satisfied.

One of the most relevant features of this possible worlds semantics lies in it provides an *internal* perspective of Kripke models. As we state that a sentence such as $\Diamond\varphi$ is satisfied in $\mathfrak{M}$ at w if and only if there exists a w′ in $\mathfrak{M}$ which is accessible from it and where φ is satisfied, what we are doing is to valuate that formula inside a model and regarding one (or some) possible world. That is why Modal Logic semantics is internal, for we valuate formulae at a specific point of a certain model.

Following Patrick Blackburn´s metaphor (2006, 332), we may conceive a modal formula like a creature located inside a model at a certain point which is compelled to change its position on the basis of some "transition rules" established by the accessibility relation. And that means that, in some sense, modal formulae truth values are context-dependents, for they depend on the possible worlds in which they are valuated.

First-Order Logic (FOL), on the contrary, does not provide an internal perspective of models but an *external* one. Indeed, in FOL, formulae valuation does not answer to some points inside a model, that is, their truth values do not depend on any kind of contextual information. Formulae are plainly true or false in a structure.

This distinction between internal/external perspective is thus closely related to intensional/extensional one. Classical Logic is extensional because formulae are true or false. But Modal Logic is intensional because formulae are true or false *according to a set of possible worlds of a model*. Formally speaking, in Modal Logic we say that a formula φ is satisfied at a world w of a model $\mathfrak{M}$, in symbols $\mathfrak{M}, w \vDash \varphi$, whereas in First-Order Logic we say that a formula φ is true in a structure or, *mutatis mutandis*, in a model $\mathfrak{M}$: $\mathfrak{M} \vDash \varphi$.

Due to this difference between perspectives Prior opted by Modal Logic instead of First-Order one to build Temporal Logic. His main idea is that our language and our thought possess an internal perspective of time by which it is represented by fixing a past moment and a future one in relation to a changing *now*. A logic of time must observe such internal perspective.

As Modal Logic (ML) is the best option for reflecting it (due to the reasons we have already seen), Prior´s proposal consists in extending ML vocabulary with new temporal operators which allow to represent sentences such as «It will be p sometime in the future» or «It has always been p in the past». Formulae resulting from these operators are valuated like modal ones, that is, by means of possible worlds semantic. Although adapted to moments of time now, and by doing so we can express things as «I will be rich sometime in the future» (1), which will be true at this moment (t_0) if and only if there exists at least a future moment t_1 (whenever this may be) when I be rich.

The problem of this approach is however clear: even though thanks to TL it is possible to formalize and valuate formulae referred to temporal events, the expressivity of such formulae is restricted for they cannot allude, e.g., to specific instants. We are not able to express something as «I will be rich on 15th May» (2) or «I am currently rich» (3). To do so we have to appeal to FOL mechanisms.

Quantification, the use of non-logical constants (individual constants, functors and predicates) and the equality relation are some of the greatest advantages which FOL possesses against ML or TL, as they allow to make reference to particular points inside a structure (model). The result of combining the internal perspective of Modal Logic with the external perspective of First-Order one is Hybrid Logic (HL), which as well as TL was developed by Prior.

HL novelty lies in it extends ML language (and consequently TL one too) with new operators and with a new sort of propositional symbol, namely, nominals, which can be merged with other formulae to compose more complex ones and that allow to *name* a specific point inside a model stating the formula they bind is true at that point, and only at that.

This paper aims to expose how HL came up as an ML and TL extension, and to explain why Hybrid Logic is very useful to, among other things, represent temporal information. In order to do so we shall follow the very same path we have followed in this introduction: in section 2 we will introduce the basic features of Modal Logic and the minimal temporal logic developed by Prior; in section 3 we will set the difference between these two systems and First-Order Logic in terms of internal/external perspectives; in section 4 we will explain how Hybrid Logic is developed from combining both perspectives and what its novelty in relation with ML and TL is; in section 5 we will state the conclusions of this paper and all the work which remains to be done in the field of Hybrid Logic; section 6 will be composed by an appendix where we will define the language we are going to use in this paper as well as by an abbreviation index; and finally in section 7 references are specified.

2. Modal Logic and Temporal Logic

Hybrid Logic is largely an extension of Modal Logic and it also comes up due to Arthur Prior and his studies in Temporal Logic. That is why it is essential to know the most relevant aspects of these two systems to fully understand what Hybrid Logic constitutes. This is going to be the aim of this section. We will split it in two parts: in the first one we will briefly explain Modal Logic syntax and semantics; in the second one we will present the minimal temporal logic devised by Prior from a contemporary point of view.

2.1. Basic Modal Logic

Classical Propositional Logic alphabet is composed by $\mathcal{L}$:

$$\mathrm{p} \mid \neg\varphi \mid \varphi \wedge \psi \mid \varphi \rightarrow \psi,$$

where $\mathrm{p} \in \mathrm{PROP}$, $\varphi, \psi \in \mathrm{WFF}$ (see appendix 6.1) and the rest of connectives ($\vee$ and $\equiv$) can be defined from $\neg$ and $\wedge$ (or from $\neg$ and $\rightarrow$).

On the other hand, its semantics is based on the notion of *interpretation*, which is a function V assigning a truth value True (1) or False (0) to every formula. V is recursively defined from f, which is a function assigning a truth value to every propositional variable, i.e., f: PROP $\rightarrow \{1, 0\}$.

V fulfills the following conditions, for every $\mathrm{p} \in$ PROP and every $\varphi, \psi \in$ WFF:

1. $\mathrm{V}(\mathrm{p}) = f(\mathrm{p})$,
2. $\mathrm{V}(\neg\varphi) = 1$ iff $\mathrm{V}(\varphi) = 0$,
3. $\mathrm{V}(\varphi \wedge \psi) = 1$ iff $\mathrm{V}(\varphi) = 1$ and $\mathrm{V}(\psi) = 1$,
4. $\mathrm{V}(\varphi \rightarrow \psi) = 1$ iff $\mathrm{V}(\varphi) = 0$ or $\mathrm{V}(\psi) = 1$.

Consequence relation is thus defined in relation with V, so a formula φ is a consequence of a set Γ of formulae if and only if there is no interpretation where every element of Γ is true and φ is not. We will symbolize it as

$$\Gamma \vDash \varphi,$$

where $\vDash$ stands for the semantic consequence relation.

In the case of Modal Logic this syntax is extended with new operators. Let $\mathcal{L}_M$ be ML language. $\mathcal{L}_M$ is composed by (Blackburn and van Benthem, 2007, 3):

- The PROP set.
- The MOD set of modal operators, where MOD = {m, *m'*, *m''*, ...} and there is an accessibility relation for each m.

The distinct elements of MOD can represent any modal operator. We are usually used to □ and ◇ (both related to the same $\mathrm{m} \in$ MOD). But they can also represent to K and B, which are the epistemic operators for knowledge and belief respectively; to P, H, F and G, which as we have seen are the temporal operators used in TL; or to many others. However, in this part they will only stand for □ and ◇.

$\mathcal{L}_M$ is thus formally defined as:

$$\mathrm{p} \mid \neg\varphi \mid \varphi \wedge \psi \mid \varphi \rightarrow \psi \mid \Diamond\varphi \mid \Box\varphi.$$

Any combination of these formulae constitutes a well-formed formula of $\mathcal{L}_M$. It belongs to the set WFF of well-formed formulae of $\mathcal{L}_M$. The proposition

$$\Diamond (\mathrm{r} \wedge \mathrm{p}) \wedge \Diamond (\mathrm{r} \wedge \mathrm{q}) \rightarrow \Diamond (\mathrm{p} \wedge \mathrm{q}) \qquad (4)$$

would be an instance such formulae. Besides, □ and ◇ are interdefinable: $\Box \equiv \neg\Diamond\neg$ and $\Diamond \equiv \neg\Box\neg$.

The alphabet of basic Modal Logic is therefore just the same as Classical Logic but with the addition of $\Box$ and $\Diamond$, so from a syntactic point of view they are not very different. The main difference is on semantics.

As we have pointed out Modal Logic is based on possible worlds semantics, which responds to the so-called *Kripke frames (or models)*. A Kripke frame is a tuple

$$\mathfrak{F} = \langle W, R \rangle,$$

where $W \neq \emptyset$ is the set of possible worlds and $R \subseteq W \times W$ is the accessibility relation between them.

From $\mathfrak{F}$ we can build a model by adding a valuation function. Let V_M be such function. What V_M makes is to assign subsets of W to propositional variables in such a way that the set of possible worlds assigned to a variable e.g. p is the set of worlds where that variable is true. Formally speaking, V_M: PROP $\rightarrow \wp(W)$. The result of adding V_M to $\mathfrak{F}$ is the model $\langle W, R, V_M \rangle$, denoted by $\mathfrak{M}_K$.

In order to express "φ is satisfied at a world w of model $\mathfrak{M}_K$" we write $\mathfrak{M}_K, w \vDash \varphi$. The satisfiability conditions of the rest of formulae are defined as follows, for any $\varphi, \psi \in$ WFF; w, w' $\in W$ and p $\in$ PROP:

1. $\mathfrak{M}_K, w \vDash p$ iff $w \in V_M(p)$,
2. $\mathfrak{M}_K, w \vDash \neg\varphi$ iff $\mathfrak{M}_K, w \nvDash \varphi$,
3. $\mathfrak{M}_K, w \vDash \varphi \wedge \psi$ iff $\mathfrak{M}_K, w \vDash \varphi$ and $\mathfrak{M}_K, w \vDash \psi$,
4. $\mathfrak{M}_K, w \vDash \varphi \rightarrow \psi$ iff $\mathfrak{M}_K, w \nvDash \varphi$ or $\mathfrak{M}_K, w \vDash \psi$,
5. $\mathfrak{M}_K, w \vDash \Diamond\varphi$ iff $\exists w' \in W$ such that wRw' and $\mathfrak{M}_K, w' \vDash \varphi$,
6. $\mathfrak{M}_K, w \vDash \Box\varphi$ iff $\forall w' \in W$ if wRw' then $\mathfrak{M}_K, w' \vDash \varphi$.

A formula is globally satisfied in a model $\mathfrak{M}_K$ if it is satisfied at every world of that model. We will symbolize it as $\mathfrak{M}_K \vDash \varphi$. A formula is valid if it is globally satisfied in every model. $\vDash \varphi$ symbolizes such thing. Finally, a formula φ is consequence of a set Γ of formulae if for every model $\mathfrak{M}_K$, every world w of $\mathfrak{M}_K$ and every $\gamma \in \Gamma$, if $\mathfrak{M}_K, w \vDash \gamma$ then $\mathfrak{M}_K, w \vDash \varphi$. And as well as in Modal Logic $\Gamma \vDash \varphi$ denotes such thing.

This system we have introduced is the most basic one of ML, and it is said it is minimal because it does not make any particular requirement from accessibility relation, i.e., it does not impose any property (reflexivity, transitivity, symmetry, etc.) when dealing with possible worlds. We will call this system K. Hence we have denoted Modal Logic models by $\mathfrak{M}_K$.

The valuation of a formula such as (4) in K would carry out as follows:

$$\mathfrak{M}_K, w \vDash \Diamond(r \wedge p) \wedge \Diamond(r \wedge q) \rightarrow \Diamond(p \wedge q)$$

holds if and only if it is verified that if

$$\mathfrak{M}_K, w \vDash \Diamond(r \wedge p) \tag{4.1}$$

and

$$\mathfrak{M}_K, w \vDash \Diamond(r \wedge q) \tag{4.2}$$

then

$$\mathfrak{M}_K, w \vDash \Diamond(p \wedge q). \tag{4.3}$$

(4.1) holds if there exists at least a possible world w' accessible from w where (r ∧ p) holds. (4.2) holds if there exists at least a possible world w'' accessible from w where (r ∧ q) holds. And (4.3) holds if there exists a possible world accessible from w where (p ∧ q) holds. As p is true at world w' and q does it at world w'' there is no guarantee of p and q being true at the same world accessible from w, so (4.3) is false. That then means that (4) is false too, for its antecedent is true but its consequent it is not. Therefore, formally,

$$\mathfrak{M}_K, w' \vDash r \wedge p,$$
$$\mathfrak{M}_K, w'' \vDash r \wedge q$$

and then

$$\mathfrak{M}_K, w \nvDash \Diamond(p \wedge q).$$

In consequence, in a model where $R = \{\langle w, w'\rangle, \langle w, w''\rangle\}$

$$\mathfrak{M}_K, w \nvDash \Diamond(r \wedge p) \wedge \Diamond(r \wedge q) \rightarrow \Diamond(p \wedge q).$$

In other words, given the model $\mathfrak{M}_K$ of K where R has any particular property, if $R = \{\langle w, w'\rangle, \langle w, w''\rangle\}$ and $w'' \notin V_M(p)$, $w' \notin V_M(q)$ and $w', w'' \in V_M(r)$ then, if $\mathfrak{M}_K, w' \vDash r \wedge p$ and $\mathfrak{M}_K, w'' \vDash r \wedge q$, $\mathfrak{M}_K, w \nvDash \Diamond(p \wedge q)$ and thus $\mathfrak{M}_K, w \nvDash \Diamond(r \wedge p) \wedge \Diamond(r \wedge q) \rightarrow \Diamond(p \wedge q)$.

Let us look at the valuation conditions of V both in Classical Propositional Logic and in Modal Logic. As we can see, in the second one V_M depends on the model and on the set of possible worlds of that model. In the first one this does not happen, however, and it shows that whereas ML formulae are intensionally interpreted Classical Propositional Logic ones are extensionally interpreted.

It is this intensionality the one which highlights the internal perspective of Modal Logic that we spoke about at the Introduction and that we shall closely discuss in Section 3. By now it is enough to know Prior´s temporal logic is one of the systems of Modal Logic that better captures that perspective for to him, as we exist in time and we make use of it in a context-dependent way by means of our utterances, a logic which tries to reflect temporal information may respect such dependence. Hence he appealed Modal Logic to do so.

2.2. Prior´s Temporal Logic

Prior´s main motivation for developing Temporal Logic was eminently philosophical. In 1941 John Findlay published "Time: A Treatment of Some Puzzles", an article where he exposes some problems derived from our conception of time and change. From then on Prior set out to build a formal system similar to Modal Logic that allows to mirror the internal perspective of time that, according to him, our language and out thought possess. His aim was anything but to formally solve those problems, specially determinism.

Prior mainly developed TL in three papers: the article "Diodoran Modalities" (1955) and the books *Time and Modality* (1962) and *Past, Present and Future* (1967). It is on them in which we shall base this section on. Although we will highlight (1967) for it is on it where the minimal system of Temporal Logic is definitely presented (Prior 1967, 176). Our way of exposing it will nevertheless differ from his as Prior´s notation is Polish one.

If K is the basic system of Modal Logic we have presented above, let K_T be the system (also basic) of Temporal Logic we are going to present hereunder.

K_T language, $\mathcal{L}_{KT}$, is composed by $\mathcal{L}$ plus the temporal operators F, P, G and H. F can be read as "it will be sometime in the *F*uture that...", P can be read as "it was sometime in the *P*ast that...", G can be read as "it will always *G*oing to be in the future..." and H can be read as "it *H*as always been in the past that...". $\mathcal{L}_{KT}$ is consequently defined as:

$$p \mid \neg\varphi \mid \varphi \wedge \psi \mid \varphi \rightarrow \psi \mid F\varphi \mid P\varphi \mid G\varphi \mid H\varphi.$$

And as before, any combination of such kind of formulae constitutes a well-formed formula of $\mathcal{L}_{KT}$. That is, it belongs to the set WFF of well-formed formulae of $\mathcal{L}_{KT}$. A proposition such as (4) (page 38) may be now arose in terms of

$$F(r \wedge p) \wedge F(r \wedge q) \rightarrow F(p \wedge q) \qquad (5)$$

or of

$$P(r \wedge p) \wedge P(r \wedge q) \rightarrow P(p \wedge q) \qquad (6)$$

As we are going to see in a moment at the conditions of the valuation function, F and P are similar to Modal Logic possibility operator, whereas G and H are with regard to necessity one. And just like □ and ◇ are interdefinable, F and G and P and H do are too: $F \equiv \neg P \neg$, $P \equiv \neg F \neg$, $G \equiv \neg H \neg$ and $H \equiv \neg G \neg$.

Regarding to its semantics Temporal Logic, being modal, is also based on possible worlds Kripke semantics. However, instead of $\langle W, R \rangle$ its frames will consist on pairs such as $\langle T, < \rangle$ where T is the nonempty set of temporal instants and < is the accessibility relation between them but understood as an ulteriority relation. So, $T = \{t, t', t'', ...\}$[1] and $< \subseteq T \times T$.

From $\langle T, < \rangle$ a model, $\mathfrak{M}_{KT}$, is defined as a pair $\langle T, <, V_T \rangle$ where V_T is evidently the interpretation function. Like V_M, V_T aims to assign a subset of temporal instants of T to each formula: the set where that formula is true. In formal terms, V_T: WFF $\rightarrow \wp(T)$.

As usual, we shall denote the satisfaction of a formula φ in the model $\mathfrak{M}_{KT}$ and at instant t as $\mathfrak{M}_{KT}, t \vDash \varphi$. Satisfaction conditions, for whatever $\varphi, \psi \in$ WFF, t, $t' \in T$ and propositional variable p, are:

1. $\mathfrak{M}_{KT}, t \vDash p$ iff $t \in V_T(p)$,
2. $\mathfrak{M}_{KT}, t \vDash \neg\varphi$ iff $\mathfrak{M}_{KT}, t \nvDash \varphi$,
3. $\mathfrak{M}_{KT}, t \vDash \varphi \wedge \psi$ iff $\mathfrak{M}_{KT}, t \vDash \varphi$ and $\mathfrak{M}_{KT}, t \vDash \psi$,
4. $\mathfrak{M}_{KT}, t \vDash \varphi \rightarrow \psi$ iff $\mathfrak{M}_{KT}, t \nvDash \varphi$ or $\mathfrak{M}_{KT}, t \vDash \psi$,
5. $\mathfrak{M}_{KT}, t \vDash F\varphi$ iff $\exists t' \in T$ such that $t < t'$ and $\mathfrak{M}_{KT}, t' \vDash \varphi$,
6. $\mathfrak{M}_{KT}, t \vDash P\varphi$ iff $\exists t' \in T$ such that $t' < t$ and $\mathfrak{M}_{KT}, t' \vDash \varphi$,
7. $\mathfrak{M}_{KT}, t \vDash G\varphi$ iff $\forall t' \in T$ if $t < t'$ then $\mathfrak{M}_{KT}, t' \vDash \varphi$,
8. $\mathfrak{M}_{KT}, t \vDash H\varphi$ iff $\forall t' \in T$ if $t' < t$ then $\mathfrak{M}_{KT}, t' \vDash \varphi$.

1. stands for propositional variables truth conditions, which are true at instant t of model $\mathfrak{M}_{KT}$ if and only if t belongs to the set of temporal instants where that variable is true. 2. represents the principle of bivalence according to which if a formula is true then its negation must be false, and vice versa. 3. and 4. reflect the interpretation of conjunction and conditional. 5. shows the truth conditions of a formula of the kind $F\varphi$, which is satisfied at instant t of model $\mathfrak{M}_{KT}$ if and only if there exists an instant t' after it where φ is satisfied. 6. shows the same but from a formula as $P\varphi$, which is satisfied at instant t of model $\mathfrak{M}_{KT}$ if and only if there exists an instant t' before it where φ is satisfied. 7., on the other hand, sets that a formula of the kind $G\varphi$ is satisfied at instant t of model $\mathfrak{M}_{KT}$ if and only if φ is satisfied at every instant after t of that model. Finally, 8. states that a formula as $H\varphi$ is

[1] In Temporal Logic, instants are usually represented by means of variable t along with integer numbers as subscripts. The current moment, where the valuation is carried out, is always symbolized as t_0. For further moments, the t subscript value is a positive integer, and for previous moments its value is a negative one.

satisfied at instant t of model $\mathfrak{M}_{KT}$ if and only if φ is satisfied at every instant before t of that model.

Global satisfaction, validity and semantic consequence relation are defined just as in ML, although in this case adapted to the model $\langle T, <, V_T \rangle$. Besides, $\mathfrak{M}_{KT}$ does not make any demand to < either.

Let us see how a formula like (4) would be valuated in K_T, but on its (5) and (6) versions. For (5) being true at instant t of model $\mathfrak{M}_{KT}$, that is,

$$\mathfrak{M}_{KT}, t \vDash F(r \wedge p) \wedge F(r \wedge q) \rightarrow F(p \wedge q),$$

there must be held that if $F(r \wedge p)$ and $F(r \wedge q)$ are true then $F(p \wedge q)$ has to be so. If $F(r \wedge p)$ is true at t means there exists another instant t' after it where $r \wedge p$ is true. The same applies to $F(r \wedge q)$: there must be an instant t'' after t where $r \wedge q$ is true. As in (4), that does not necessarily mean that there exists an instant later than t where $p \wedge q$ is true and therefore $F(p \wedge q)$ is false at t in some models.

What follows from this is that (5) is false, for its antecedent is true but its consequent is not, and then

$$\mathfrak{M}_{KT}, t \nvDash F(r \wedge p) \wedge F(r \wedge q) \rightarrow F(p \wedge q).$$

That is to say, if $< = \{\langle t, t' \rangle, \langle t, t'' \rangle\}$ and $t' \notin V_T(q)$ or $t'' \notin V_T(p)$ then it happens that

$$\mathfrak{M}_{KT}, t' \vDash r \wedge p,$$
$$\mathfrak{M}_{KT}, t'' \vDash r \wedge q.$$

But

$$\mathfrak{M}_{KT}, t \nvDash F(p \wedge q)$$

and so

$$\mathfrak{M}_{KT}, t \nvDash F(r \wedge p) \wedge F(r \wedge q) \rightarrow F(p \wedge q).$$

In the case of (6) exactly the same takes place, but regarding to instants t' and t'' prior to t: if $< = \{\{\langle t', t \rangle, \langle t'', t \rangle\}$ and $t' \in V_T(p)$, $t'' \in V_T(q)$ and $t', t'' \in V_T(r)$ then, if $\mathfrak{M}_{KT}, t' \vDash r \wedge p$ and $\mathfrak{M}_{KT}, t'' \vDash r \wedge q$, $\mathfrak{M}_{KT}, t \nvDash P(r \wedge p) \wedge P(r \wedge q) \rightarrow P(p \wedge q)$ holds.

Nonetheless, natural language does not work like that. For instance, let us suppose we are talking to a friend about what we are going to do this summer and we are telling him/her we are going to go to a water park where we will sunbath while we drink beer at its slow river–fun guaranteed. If we were to speak logically, we may say something as: «If I will go sometime in the future (this simmer) to a water park to enjoy sunbathing and I

will go sometime in the future to a water park to drink beer at its slow river then I will go sometime in the future to sunbath and drink beer at a slow river».

If variable r stands for "going to a water park", variable p stands for "sunbath" and variable q stands for "drinking a beer at a slow river" then $F(r \wedge p) \wedge F(r \wedge q) \rightarrow F(p \wedge q)$, which is (5), stands for the aforementioned proposition. A proposition that is true today (t) as long as it is true that this summer (t′) we will go to a water park where we will sunbath and drink a beer at its slow river. But it would be false if we do not set that the moment where p and q are true is the same.

In any case, what this example shows is that thanks to Modal Logic mechanisms Temporal Logic is able to reflect both time flows and that internal perspective of temporality we talked about at the beginning and that, according to Prior, our language possesses. Indeed, whenever we allude to some fact, we do so by taking as point of reference the present moment and by setting that such fact either happened at some earlier moment or it is happening now or it will happen at a later moment. That context-dependence is the main characteristic of Modal Logic, as we have said, and hence TL constitutes a nice example about how ML deals with models from inside.

However, if we look at conditions 5 and 6 of V_M in ML and at conditions 5-8 of V_T in TL we shall see that we may use First Order Logic to capture at least a part of that internal perspective of modal representation. ML (and thus TL) then is weaker than FOL, and in fact by means of what is known as *Standard Translation* Modal Logic can be represented by First Order Correspondence Language mechanisms. In particular, by using free variables to contextualize FOL formulae valuations based on certain individuals (counterparts of possible worlds).

Therefore, even although First Order Logic provides an external point of view of models it can be employed to reflect Modal Logic internal perspective, and that is what we are going to present in the next section.

3. Standard Translation and Modal Logic Problems

At the beginning of section 2 we said that Hybrid Logic maintains a close relation with Modal Logic. But it is also closely related to First Order Logic. Specifically, to a particular part of FOL called *First Order Correspondence Language* (FOCL)[2]. ML and FOCL in turn possess a very interesting connection, which is necessary to know to understand what the true innovation of Hybrid Logic is and what it gains with respect to ML, TL and FOCL.

The aim of this section is thus to show that connection between Modal Logic and First Order Correspondence Language as well as the problems of the former in representing certain kind of propositions. To do so the section will be divided in two parts: on the first

[2]From now on any allusion to FOL must be understood as referred to this language and not to First Order Logic entirely.

one we will address the relation between ML and FOCL, and on the second one we will deal with ML problems.

3.1. Relation between Modal Logic and First Order Correspondence Language

Modal Logic formulae may be expressed by First Order Correspondence Language formulae with at least one free variable. Nevertheless, this does not happen conversely: First Order Correspondence Language formulae cannot be expressed by Modal Logic ones for the latter is weaker than the former. There are FOCL formulae which are not expressible in ML. That is why it is possible to conceive ML as a part of its corresponding first order language.

Modal Logic Kripke models are no (Blackburn and van Benthem 2007, 10) more than relational structures composed by a domain over which quantification is carrying out (W in the case of ML and T in the case of TL), a range of binary relations over that domain (R in the case of ML and < in the case of TL) and another range of unary relations which are applied to formulae of that structure (V_M in the case of ML and V_T in the case of TL).

The problem —one may think— is, if so, then we do not actually have to resort to Modal Logic to talk about possible worlds. It is enough to First Order Logic to do so. Indeed, if we add to FOL language a binary relator R applied to every element of MOD, i.e., a relation R^m for every $m \in MOD$, and a unary relation P for every element of PROP, i.e., a relation P for every $p \in PROP$, then we can set a correspondence between ML (or TL) formulae and FOL formulae by means of the Standard Translation. That is what FOCL consists on.

Let ST_x be a function assigning to every modal formula its corresponding first order formula. ST_x consists on the following, for every propositional variable p, whatever $\varphi, \psi \in WFF$ and whatever [m], $\langle m \rangle \in MOD$ (where [m] represents any modal operator similar to □, that is, whose valuation depends on every possible world accessible from the actual one, and ⟨m⟩ represents any modal operator similar to ◇, that is, whose valuation depends on some possible world accessible from the actual one) (Blackburn 2006, 334):

1. $ST_x(p) = P(x)$,
2. $ST_x(\neg\varphi) = \neg ST_x(\varphi)$,
3. $ST_x(\varphi \wedge \psi) = ST_x(\varphi) \wedge ST_x(\psi)$,
4. $ST_x(\varphi \rightarrow \psi) = ST_x(\varphi) \rightarrow ST_x(\psi)$,
5. $ST_x(\langle m \rangle \varphi) = \exists y\,(R^m\langle x, y\rangle \wedge ST_y(\varphi))$,
6. $ST_x([m]\varphi) = \forall y\,(R^m\langle x, y\rangle \rightarrow ST_y(\varphi))$.

1-6 are very similar to 1-6 V_M conditions we saw at part 2.1. But what they set is a correspondence between Modal Logic and First Order Logic by resorting to free variable *x*, whose importance we shall see hereunder.

1. shows that a ML propositional variable may be translated via FOCL as a predicate P whose argument is *x*. In Modal Logic we saw that a proposition such as p is satisfied at world w of model $\mathfrak{M}_K$ if and only if w belongs to the set of possible worlds where p is true. That is what $\mathfrak{M}_K$, w ⊨ p iff w ∈ V_M(p) meant.

As in First Order Logic formulae valuations do not depend on Kripke semantics we cannot assert the same while we are talking about p truth. To do so we use free variables (*x* in this case), which are what reflect the internal perspective of Modal Logic through Classical Logic. By assigning a value to *x* what we are doing is something similar to what happens in ML when we set that p is true at w, namely, we indicate that the formula concerned by *x* is true (or false) in *x*. P(*x*) then means that every propositional symbol p is true in *x*. $\mathfrak{M}_K$, w ⊨ p iff w ∈ V_M(p) and P(*x*) are therefore analogous, for assigning a value to *x* is equivalent to valuate a modal formula at some world of a model.

Conditions 3. and 4. work similarly so let us focus on 5. and 6. ones. As we have a formula as $\Diamond\varphi$ in Modal Logic we say it is satisfied at world w of model $\mathfrak{M}_K$ if and only if there exists some world w' accessible from w where φ is true. $\mathfrak{M}_K$, w ⊨ $\Diamond\varphi$ iff ∃w' ∈ W such that wRw' and $\mathfrak{M}_K$, w' ⊨ φ sets such thing. In the case of Classical Logic the valuation of $\Diamond\varphi$ does not differ too much. The only remarkable difference is naturally that the truth of φ does not depend on any possible world of $\mathfrak{M}_K$ but on free variable *x* and bound variable *y*.

So, what condition 5. shows is that a formula as ⟨m⟩φ is true in Classical Logic, on the basis of *x*, if and only if there exists at least one *y* such that *x* and *y* are related and φ is true in *y*. Something similar happens to formulae of the kind [m]φ, which are true in Classical Logic on the basis of *x* if and only if, for every variable *y*, if *x* and *y* are related then φ is true in *y*.

Consequently,

$$\mathfrak{M}_K, w \vDash \Diamond\varphi \text{ iff } \exists w' \in W \text{ such that } wRw' \text{ and } \mathfrak{M}_K, w' \vDash \varphi$$

and

$$ST_x(\langle m\rangle\varphi) = \exists y\,(R^m\langle x, y\rangle \wedge ST_y(\varphi)),$$

are equivalent. And

$$\mathfrak{M}_K, w \vDash \Box\varphi \text{ iff } \forall w' \in W \text{ if } wRw' \text{ then } \mathfrak{M}_K, w' \vDash \varphi$$

and

$$ST_x([m]\varphi) = \forall y\,(R^m\langle x, y\rangle \rightarrow ST_y(\varphi)),$$

are too. This evinces that modal formulae and their first order translation express the same.

Standard Translation may be also applied to Temporal Logic. There would just have to modify conditions 5. and 6. to be adapted to F, P, G and H operators. In order to do so we need to turn the order of P and H R^m pairs around to reflect conditions 6. and 8. of V_T:

5*. $ST_x(F\varphi) = \exists y\,(R^m\langle x, y\rangle \wedge ST_y(\varphi))$,

6*. $ST_x(P\varphi) = \exists y\,(R^m\langle y, x\rangle \wedge ST_y(\varphi))$,

7*. $ST_x(G\varphi) = \forall y\,(R^m\langle x, y\rangle \rightarrow ST_y(\varphi))$,

8*. $ST_x(H\varphi) = \forall y\,(R^m\langle y, x\rangle \rightarrow ST_y(\varphi))$.

An example of how Standard Translation would work is the following one, where R is a unary relation for every $r \in PROP$ and Q is a unary relation for every $q \in PROP$; and P keeps the previous meaning:

$$\Diamond(r \wedge p) \wedge \Diamond(r \wedge q) \rightarrow \Diamond(p \wedge q) =$$

$$ST_x(\Diamond(r \wedge p) \wedge \Diamond(r \wedge q) \rightarrow \Diamond(p \wedge q)) =$$

$$\exists y\,(R^m\langle x, y\rangle \wedge (R(y) \wedge P(y))) \wedge \exists y\,(R^m\langle x, y\rangle \wedge (R(y) \wedge Q(y))) \rightarrow \exists y\,(R^m\langle x, y\rangle \wedge (P(y) \wedge Q(y))).$$

The formula we have recursively translated into Classical Logic is our example (4), whose Standard Translation is $ST_x(\Diamond(r \wedge p) \wedge \Diamond(r \wedge q) \rightarrow \Diamond(p \wedge q))$. By condition 4. of ST_x, $ST_x(\Diamond(r \wedge p) \wedge \Diamond(r \wedge q) \rightarrow \Diamond(p \wedge q))$ is equal to $ST_x(\Diamond(r \wedge p) \wedge \Diamond(r \wedge q)) \rightarrow ST_x(\Diamond(p \wedge q))$. And by the recursive application of conditions 1., 3. and 5. we obtain $\exists y\,(R^m\langle x, y\rangle \wedge (R(y) \wedge P(y))) \wedge \exists y\,(R^m\langle x, y\rangle \wedge (R(y) \wedge Q(y))) \rightarrow \exists y\,(R^m\langle x, y\rangle \wedge (P(y) \wedge Q(y)))$, which is the full Standard Translation of (4).

(5) and (6) translation would be similar, but by following conditions 5*.-8*., instead of 5. and 6., this time. (5) would then be:

$$F(r \wedge p) \wedge F(r \wedge q) \rightarrow F(p \wedge q) =$$

$$ST_x(F(r \wedge p) \wedge F(r \wedge q) \rightarrow F(p \wedge q)) =$$

$$\exists y\,(R^m\langle x, y\rangle \wedge (R(y) \wedge P(y))) \wedge \exists y\,(R^m\langle x, y\rangle \wedge (R(y) \wedge Q(y))) \rightarrow \exists y\,(R^m\langle x, y\rangle \wedge (P(y) \wedge Q(y))).$$

(6), on its part, would be alike but with R^m applied to $\langle y, x\rangle$.

The most interesting thing about Standard Translation is it preserves satisfiability, i.e., both

$$\Diamond(r \wedge p) \wedge \Diamond(r \wedge q) \rightarrow \Diamond(p \wedge q) \text{ and } ST_x(\Diamond(r \wedge p) \wedge \Diamond(r \wedge q) \rightarrow \Diamond(p \wedge q))$$

and

$$F(r \wedge p) \wedge F(r \wedge q) \rightarrow F(p \wedge q) \text{ and } ST_x(F(r \wedge p) \wedge F(r \wedge q) \rightarrow F(p \wedge q))$$

(along with their analogous with (6)) are equisatisfiable, what means that if one of them are satisfiable in First Order Logic then its corresponding one will be so. We shall denote such property by means of symbol ≈. Hence, we picture that

$$\Diamond(r \wedge p) \wedge \Diamond(r \wedge q) \rightarrow \Diamond(p \wedge q)$$

and

$$ST_x(\Diamond(r \wedge p) \wedge \Diamond(r \wedge q) \rightarrow \Diamond(p \wedge q))$$

are equisatisfiable as

$$\Diamond(r \wedge p) \wedge \Diamond(r \wedge q) \rightarrow \Diamond(p \wedge q) \approx ST_x(\Diamond(r \wedge p) \wedge \Diamond(r \wedge q) \rightarrow \Diamond(p \wedge q)).$$

Such result may be expressed by the following proposition (Blackburn and van Benthem 2007, 11; Blackburn 2006, 335):

> **Definition 3.1.1. ML-FOCL Equisatisfiability (ML ≈ FOCL)** For any basic modal formula φ, any model $\mathfrak{M}_K$ and any possible world w in $\mathfrak{M}_K$, $\mathfrak{M}_K$, $w \vDash \varphi$ if and only if $\mathfrak{M}_K \vDash ST_x(\varphi)\ [x \leftarrow w]$.

$[x \leftarrow w]$ means that x takes world w value, that is, w is assigned to free variable x. The proof of **ML ≈ FOCL** directly arises from conditions 1.-6. of Standard Translation. Every modal formula thus can be transformed into its corresponding First Order Logic one but, as we remarked at the beginning of this section, not every FOCL formula may be transformed into a ML one. $\neg(xRx)$ or $\neg(x < x)$ would be two examples. Indeed, whereas in FOL we can express properties such as irreflexivity by means of $\neg(xRx)$ or $\neg(x < x)$, in ML we cannot. The reason lies in any modal formula is satisfied at every irreflexive points of a model. To explain why it is necessary to introduce the notion of *bisimulation*.

Let $\mathfrak{M}_K$ and $\mathfrak{M}_K$* be two whatever models of ML and let B be a binary relation between these two models. B is a bisimulation between $\mathfrak{M}_K$ and $\mathfrak{M}_K$* if the following holds:

- If B is applied to two worlds w and w′ of $\mathfrak{M}_K$ and $\mathfrak{M}_K$* respectively then, for every propositional variable p, if w is on the set of possible worlds of $\mathfrak{M}_K$ where p is true, w′ has also to be on the set of possible worlds of $\mathfrak{M}_K$* where p is true.
- If w and w′ are related through B and w is related in $\mathfrak{M}_K$ through R to another possible world v, then w′ must be related in $\mathfrak{M}_K$* through R′ to another possible world v′ too and, besides, v is related to v′ through B.
- If w and w′ are related through B and w′ is related in $\mathfrak{M}_K$* through R′ to v′ then w is also related to v through R in $\mathfrak{M}_K$ and, moreover, v and v′ are related through B.

Formally speaking:

Definition 3.1.2. ML Bisimulation If $\mathfrak{M}_K$ and $\mathfrak{M}_K$* are two models of basic Modal Logic such that $\mathfrak{M}_K = \langle W, R, V_M \rangle$ and $\mathfrak{M}_K^* = \langle W', R', V'_M \rangle$ then the relation $B \subseteq W \times W'$ is a bisimulation between $\mathfrak{M}_K$ and $\mathfrak{M}_K$* if the following conditions are satisfied:

1. If wBw′ then, for any $p \in PROP$, $w \in V_M(p)$ if and only if $w' \in V'_M(p)$.
2. If wBw′ and wRv then there exists a point v′ in $\mathfrak{M}_K$* such that w′R′v′ and vBv′.
3. If wBw′ and w′R′v′ then wRv and vBv′.

We say that a possible world w in $\mathfrak{M}_K$ is bisimilar to another possible world w′ in $\mathfrak{M}_K$* if wBw′.

Condition 1. states that two bisimilar worlds satisfy the same propositions, i.e., if p is true at w in $\mathfrak{M}_K$ it will also be true at w′ in $\mathfrak{M}_K$*. 2. and 3., on their part, set that if there are two bisimilar worlds and one of them access to another in $\mathfrak{M}_K$ or $\mathfrak{M}_K$* then that accessibility relation will also hold in the other model and these two accessed worlds will be bisimilar too. Therefore, two worlds are bisimilar if they satisfy the same formulae and access to the same worlds.

What follows from this is that two bisimilar worlds also satisfy the same modal formulae, for their satisfaction only requires the access to another world(s) and that in that world(s) such formula is satisfied. In consequence, two modal formulae are not able to distinguish between two bisimilar worlds.

The clearest example is the one we have previously highlighted: $\neg(xRx)$ or $\neg(x < x)$. Let us suppose $\mathfrak{M}_K$ is a model where every propositional variable is false at every world and $\mathfrak{M}_K$* is a model where there is only a reflexive possible world where every propositional variable is false too. In this case B relates every $\mathfrak{M}_K$ world to the only $\mathfrak{M}_K$* world (w′). By the result derived from **ML Bisimulation** every $\mathfrak{M}_K$ and $\mathfrak{M}_K$* world satisfies the same modal formulae, but that means there cannot be a modal formula true at every irreflexive worlds for, if it existed, it would be true at every $\mathfrak{M}_K$ worlds while false at w′

of $\mathfrak{M}_K$*. As this is impossible for there is a bisimulation between $\mathfrak{M}_K$ worlds and w' it follows that no modal formula satisfies something as $\neg(x\mathrm{R}x)$ or $\neg(x < x)$.

Modal Logic expressive capacity is thus less than First Order Logic one. That is the reason why every ML formula can be translated into a FOCL one but not conversely. There is not a modal formula equisatisfiable to the first order formula $\neg(x\mathrm{R}x)$. Modal Logic may consequently be conceived as a proper fragment of First Order Logic. In particular, as the fragment composed by those formulae closed under bisimulation, that is, Modal Logic is that part of First Order Logic holding every formula invariant for bisimulation.

It is said a FOL formula φ with one free variable is invariant for bisimulation if its valuation at bisimilar worlds is always the same. The fact that ML is a proper fragment of FOCL follows from this definition and the so-called van Benthem´s *Characterisation Theorem* (van Eijck 2006, 15; Blackburn 2006, 337-338; Blackburn and van Benthem 2007, 21), according to which if a first order formula φ with one free variable is invariant for bisimulation then φ is equivalent to the standard translation of its corresponding modal formula.

Why thus making use of Modal Logic instead of First Order one if the former is only a small part of the latter? For two reasons, mainly:

I. The first one is because of decidability: Modal Logic is PSPACE-decidable whereas First Order Logic is not.

II. The second one is due to *perspectivism*: Modal Logic provides an internal perspective of models whereas First Order Logic one is external.

These two reasons, along with language simplicity, justify the importance of Modal Logic. Although ML has still great problems to represent certain kind of propositions, as we will see down below.

3.2. Problems of Modal Logic

In the previous section we have seen that, despite Modal Logic is properly included in First Order Correspondence Language, it is still useful as the preceding I. and II. conditions show. Specially thanks to the second one, which was Prior´s reason for choosing ML to build Temporal Logic.

However, Modal Logic main problem lies in its expressivity. As mentioned previously, there is no modal formula which satisfies an irreflexive proposition, and that constitutes a very remarkable expressive limitation. But now we are going to focus on another limitation, namely, its inability to name points inside a model.

The main characteristic of ML and TL languages is they are composed by modal expressions ($\Box\varphi$, $\Diamond\varphi$, $\mathrm{F}\varphi$, $\mathrm{P}\varphi$, $\mathrm{G}\varphi$ or $\mathrm{H}\varphi$) which allow to relativize φ truth to a range of worlds or instants (or one in the case of $\Diamond$, F and P). But they do not allow for instance

to set that φ is true at exactly *this* world (instant). First Order Logic can do it. Indeed, by means of constants and the identity relation we are able to state that certain individual possesses such and such property, or that two individuals with x property are equal. But ML and TL do not count with these mechanisms.

What directly follows from that is Temporal Logic, being based on Modal Logic, is not able to accurately represent the natural conception of time. As we allude to facts which have happened or will happen at any time different from current one we do not usually make do with claiming that such facts have happened sometime in the past or will happen sometime in the future. We generally aim to state when they happened or will happen, and Temporal Logic cannot formally represent.

In Classical Modal Logic, either purely modal or temporal, epistemic, etc., formulae are valuated regarding a point of reference which tends to be a possible world. That point of reference (or possible world) can change and so formulae truth value can change too. If for instance we say «I am going to make tea later» it is evident that such claim shall have different truth values depending on the possible world (or the moment and history in the case of Temporal Logic) where it is valuated. Nonetheless, sometimes we want to be more precise and set the exact moment where such fact happened, happens or will happen. Following our example, maybe we want to specify that we are going to make tea at five o´clock this afternoon.

The first proposition, «I am going to make tea later», can be formalized in terms of TL as F(1)p, where F is the possibly future operator, value 1 indicates p will take place one time unit (let us put eight hours) later than the utterance moment[3] and p stands for "making tea later". But the second proposition, «I am going to make tea at five o´clock this afternoon», cannot be formalized.

It cannot be so for, to do it, it is required to introduce nominals, i.e., symbols which exactly determine when a proposition is valuated. Therefore, the point of reference which will allow to set the truth value of propositions bound with nominals is just the point such nominal designates, and just that one. In our case, it may be the 4th of July at 17:00.

Following Hans Reichenbach´s (1947) distinction between point of utterance (S), point of event (E) and point of reference (R) we can see that basic Temporal Logic only counts with the first and the second ones but not with the third one. As we valuate a formula such as F(1)p at t_0, S, which is the point of valuation, is precisely t_0 whereas E, which is the point where the fact the proposition talks about takes place, is some future moment t_1 where p is true. The same applies to P, G and H.

[3]We are dealing with Metric Temporal Logic (Prior 2010, 159-170; Øhrstrøm and Hasle 1995, 231-240).

This structure allows us to formalize things as «I am going to make tea later» or «I made tea before», but it does not allow us to formalize «I had made tea», for instance, since for this claim be valuated it is necessary to count on a point of reference such that succeeds S and precedes E. Graphically:

$$E \leftarrow R \leftarrow S.$$

By stating that at t_0 we had made tea what we are asserting is there is a previous moment t_{-1}, corresponding to R, where it is true that at some previous moment t_{-2} we made tea. That moment corresponds to E.

In the same way, if we said «I will have made tea» at t_0 what we were asserting is that it is true at some future moment t_2 (R) that in some past moment t_1 (E) it is true that we will have made tea. (Although the structure of this tense is not always this way (Reichenbach, 1947).) Graphically

$$S \leftarrow E \leftarrow R.$$

Structures like these ones are what basic Temporal Logic is not able to represent. Its syntax and semantics do not allow to allude to specific instants. That is why it is narrowed. And also that is why Prior saw the necessity of embracing some FOL mechanisms to build a more expressive system than TL. That system is Hybrid Logic.

We thus have seen that Modal Logic may be subsumed under First Order Logic by means of Standard Translation. That makes the internal perspective of the former can be expressed through the external perspective of the latter and that we are able to call Modal Logic usefulness into question. Its utility lies, however, in its capacity for talking about models from inside them. In the case of Temporal Logic that internal perspective is crucial for developing a logical system which aims to represent time natural conception. The problem, nevertheless, is TL does not achieve it for it is based on Modal Logic, whose expressive ability is limited. A solution may be to merge ML/TL and FOL mechanisms in order to increase that expressivity. The result of this merging is Hybrid Logic, to which we shall devote the following section.

4. Hybrid Logic

At the beginning of this paper we said Prior´s motivation for building both Temporal Logic and Hybrid Logic was mainly philosophical. According to him any issue related to time (and almost anything) may be logically solved and that is why, as we will see, when the time comes he has to resort to Hybrid Logic to solve one of the greatest problems of TL. Problem linked to TL expressivity, which we have set out at the previous section.

At this one what we are going to do is, on the one hand, to explain Hybrid Logic motivation to, on the other hand, introduce its basic system and extend it in accordance with Prior´s approach.

4.1. Hybrid Logic Motivation

In (1908) John McTaggart raises two ways of conceiving time, i.e., two ways of understanding the arrangement of facts in time: A series and B series. The first one consists in sorting events in terms of past, present and future. The second one consists in sorting them on the basis of ulteriority relation, that is, on the basis of the before/after relation. A series accurate reflections are F, P, G and H operators whereas B series one is the $<$ relation.

By sorting facts according to whether they take place at present, at past or at future we embrace an internal conception of time, namely, we place ourselves inside it and we reflect the chain of events from the future to the past, through the present. Temporal Logic, by arising from Modal one, picks up this internal view of time by means of F, P, G and H.

In contrast, by sorting facts according to whether they take place before or after each other we embrace an external view of time, for we place ourselves from outside it and we just reflect its course through the mere succession of events. The $<$ relation, characteristic of First Order Logic, embodies this idea very well.

That Prior has developed to a larger extent Temporal Logic and that he has conferred it a greater importance on his writings proves that, for him, A series takes priority over B series. In fact, in his opinion, the former presupposes the latter. But not only that. Besides, B series has two important problems: firstly, it does not actually represent the way we experience time for our experience is not external but internal; secondly, it entails accepting the existence of instants. Indeed, in TL quantification is carrying out over instants (as we have seen in V_T 5-8 conditions) and by claiming that "There exists an instant t such that..." or "For every instant t..." what we are doing is to commit ourselves to their existence, which is very doubtful.

This is why Prior prioritizes A series over B one, although he must prove it is possible to subsume it under A, and why he resorts to Hybrid Logic.

In "Tense Logic and the Logic of Earlier and Later" (2010) Prior avers there are four types (grades) of logical-temporal entailment:

In the first one he presents TL as some kind of abbreviation of what he calls *U-calculus*, which is no more than his version of Standard Translation. What he advocates is that formulae as Fφ or Pφ are shorthand of «There exists an instant t' later than t where φ is true» and «There exists an instant t' earlier than t where φ is true», respectively, and so temporal operators may be understood as ways of synthesizing $<$ properties. In other

words: Prior appeals to Standard Translation to show that Temporal Logic may be conceived as a Temporal First Order Logic. A series is then reduced to B one.

In the second grade temporal operators are no longer subsumed under $<$ relation. They are now at the same level. The key to understand such thing lies in how atomic propositions are treated. As we saw at 2.2 section K_T semantics demands for referring to instants in order to valuate formulae. That means a proposition as p cannot be valuated, by itself, in K_T because we do not count with the reference to an instant to carry it out. To do so we would need a moment t and stating that p is true in it, i.e., to say that $\mathfrak{M}_{KT}$, $t \vDash p$. Thus, p on its own is an uncompleted formula.

Nevertheless, what Prior claims is that p actually constitutes what Nicholas Rescher and Alasdair Urquhart have called *chronologically definite propositions* (Rescher and Urquhart 1971; Prior 2010, 118; Øhrstrøm and Hasle 1995, 218). A chronologically definite proposition is that one which implicitly alludes to an instant though it does not explicitly allude to it. This means that a formula as p is equivalent to $\mathfrak{M}_{KT}$, $t \vDash p$, that is, p is a way of stating "p is true at instant t". And what results from this is A and B series are at the same conceptual level. None of them are constrained by the other one but both are related.

In the third grade this connection between both series becomes more evident thanks to the introduction of what Prior calls *world-variables*, which are just nominals. According to Prior nominals depict the (possible) world just as it is at the very moment they allude to. That implies their nature is two-fold: on the one hand, they are indexes naming a certain moment; on the other hand, they are propositions depicting the world just as it is at that very moment.

If so, that is, if nominals are not only terms which refer to time instants but they are also propositions, then one of the greatest problems of B series, namely its commitment with instants existence, is solved. By being propositions, instants are no longer fictitious entities (Prior 1967, 188-189), and as claiming that «φ is true just at instant i» is the same that claiming «It is necessary that if i then φ», it is possible to entirely derive Temporal Logic from A series plus the necessity operator, i.e., from F, P, G, H and $\Box$.

Finally, the fourth grade is no more than an attempt to define the necessity operator in terms of temporal ones. In it Prior suggests what is known as *universal modality* and besides he reduces the entire Temporal Logic (including $\Box$) to F and P operators.

Third and fourth grades allow to subsume B series under A one, and Hybrid Logic arises due to them. The system which reflects the third one is more general than the one which reflects the fourth grade, which is more specific, but in any case both are origin of Hybrid Logic. A logic that, as Prior later acknowledges in articles such as "Quasi-Propositions and Quasi-Individuals" (Prior 2010, 213-221) or "Egocentric Logic" (Prior 2010, 223-240), may be extended to any domain and allows to consider as propositions not only predicates alluding to instants but basically any predicate. Description Logic, posed by him in the second aforementioned article, results from this way of conceiving HL.

Hybrid Logic arising is then due to Prior´s pretence of proving that A series truly reflects our conception of time and that B series may be reduced to it. And it means that we can reduce one part of First Order Logic to a purely temporal logic. A reduction which just may be carried out by Hybrid Logic mechanisms. In the following section we shall see what they are.

4.2. Two Systems of Hybrid Logic

Hybrid Logic relevance both for Temporal Logic and for Modal Logic has been evinced in the previous section. In this one what we are going to do is to formally present two systems of HL: the first one is its basic system whereas the second one is the strongest system presented by Prior himself.

HL alphabet is based on ML one but with two main differences: it adds a new sort of propositional symbols and a new set of modal operators.

Apart from PROP set of propositional variables Hybrid Logic adds a second sort of symbols to represent nominals: NOM. Those symbols are *i, j, k, l*, etc.: NOM = {*i, j, k, l*, ...}. One of the most interesting and relevant characteristics of nominals is they constitute terms, that is, they are not only variables indicating a specific moment but those variables are, by themselves, propositions. Hence, if we say «*i*» what we are asserting is, on the one hand, there is an instant called "*i*" and, on the other hand, that formula is true at the instant whose name is "*i*".

Nominals can be combined with other formulae to build more complex formulae, and such formulae are true just at the instant named by the nominal. Their function is thus to name a point inside a model and to set that the proposition bound by them is true at exactly that point. Thanks to that we are able to solve the Modal Logic expressive limitation.

Recall proposition (4):

$$\Diamond(r \wedge p) \wedge \Diamond(r \wedge q) \rightarrow \Diamond(p \wedge q).$$

(4) is not valid in K, but it does be in Hybrid Logic as we substitute r by *i*. The resulting formula would be:

$$\Diamond(i \wedge p) \wedge \Diamond(i \wedge q) \rightarrow \Diamond(p \wedge q). \qquad (4^*)$$

(4*) is always true for what it states is the world where p is true and the world where q is true are the same, namely, *i*. And hence it is true that there exists a world w' (*i*) where p and q are simultaneously true.

Regarding (5) and (6) the same happens. If these propositions are transformed in

$$F(i \wedge p) \wedge F(i \wedge q) \rightarrow F(p \wedge q) \qquad (5^*)$$

and

$$P(i \wedge p) \wedge P(i \wedge q) \rightarrow P(p \wedge q) \qquad (6^*)$$

by substituting r by *i* then the moments where p and q are true overlap, and thus the consequent of these conditionals is always true.

But nominals may also formally represent propositions which Modal/Temporal logics are not able to do such as «I had made tea», which we discussed at section 3.2. In Hybrid Logic the formula representing such proposition would be:

$$P(i \wedge Pp) \qquad (7)$$

which sets there is a moment *i* earlier than t_0 and a moment t_{-1} earlier than *i* such that p is true at it. In other words, the moment where that fact takes place is earlier than the point of reference, which in turn predates the moment where (7) is uttered. (7) structure is, as we may see,

$$E \leftarrow R \leftarrow S,$$

where *i* is R, corresponding to what Reichenbach poses in (1947) and we remarked at point 3.2[4].

So nominals increase and enhance Modal/Temporal logics expressivity by allowing to allude to specific points and they constitute the first step in Hybrid Logic building. The second one lies in the satisfaction operator @. Its function consists in binding both a nominal and a propositional variable to determine the very moment that variable is true. A formula such as $@_i\varphi$, read as "At *i*, φ", sets that φ is satisfied only and exclusively at *i*. @ is therefore a modal operator whose range is just a possible world or moment, and alike □ and ◇ it fulfills the following properties:

- Distributive:

  $$@_i(\varphi \rightarrow \psi) \rightarrow (@_i\varphi \rightarrow @_i\psi).$$

- Generalization:

  $$\text{If } \vdash\varphi \text{ then } \vdash@_i\varphi,$$

for any *i*.

- Self-duality:

  $$@_i\varphi \equiv \neg@_i\neg\varphi.$$

[4]It is worthy to highlight that, notwithstanding the tremendous link between Reichenbach´s approach and Prior´s one, the latter never thought there was any. In fact, Prior thought Reichenbach´s proposal was too messy because of his distinction between R and E (Prior 1967, 13); however using nominals implies such distinction.

One of the main advantages of satisfaction operator is it provides a modal interpretation of identity relation. Indeed, thanks to @ we may represent identity in Hybrid Logic by the formula

$$@_i j,$$

which states that, at point i, j is satisfied. As i and j are nominals that means point i is identical to point j and consequently we may express $=$ by means of $@_i j$. @ is also able to express other properties of identity for which it is necessary to resort to $=$ and First Order Logic in other systems, namely, reflexivity, symmetry, transitivity and substitution (Blackburn 2006, 343):

- Reflexivity:

  $\forall x\ (x = x)$

  is represented in HL as

  $@_i i$.

- Symmetry:

  $\forall x \forall y\ (x = y \rightarrow y = x)$

  is represented in HL as

  $@_i j \rightarrow @_j i$.

- Transitivity:

  $\forall x \forall y \forall z\ (x = y \wedge y = z \rightarrow x = z)$

  is represented in HL as

  $@_i j \wedge @_j k \rightarrow @_i k$.

- Substitution:

  $\forall x \forall y\ (x = y \wedge \Psi(x) \rightarrow \Psi(y))$

  is represented in HL as

  $@_i \varphi \wedge @_i j \rightarrow @_j \varphi$,

  where Ψ symbolizes any property.

Hybrid Logic, due to @, is thus able to modally translate FOL formulae. This is why (among other things) its expressive capacity is bigger than other systems as ML and TL one.

Basic Hybrid Logic alphabet, called L_H, consists of:

$$i \mid \mathrm{p} \mid \neg\varphi \mid \varphi \wedge \psi \mid \varphi \rightarrow \psi \mid \Box\varphi \mid \Diamond\varphi \mid @_i\varphi,$$

where $i \in$ NOM, p $\in$ PROP, $\varphi, \psi \in$ WFF and $\Box$, $\Diamond$ and @ $\in$ MOD (all of them under the same accessibility relation). (4*), (5*) and (6*) would be examples of HL well-formed formulae.

As HL is no more than ML enriched with NOM and @ it addresses possible worlds semantics too. So, from the frame $\mathfrak{F} = \langle W, R\rangle$ we build the model $\mathfrak{M}_H = \langle W, R, V_H\rangle$, where W is a nonempty set of possible worlds (or states or points), R is the accessibility relation between them such that $R \subseteq W \times W$ and V_H is the valuation function assigning subsets of W both to propositional variables and to nominals, i.e., V_H: PROP $\cup$ NOM $\rightarrow \wp(W)$. Hence, V_H domain is PROP $\cup$ NOM and its range is $\wp(W)$.

Regarding nominals V_H is always a singleton subset of W, that is, for any $i \in$ NOM and any w $\in$ W, $w \in V_H(i) \equiv w = i$. In other words: every nominal has only and the same value at W. $V_H(i)$ single world is the denotation of i.

As well as in Modal Logic we shall symbolize that a formula φ is satisfied at world w of model $\mathfrak{M}_H$ as $\mathfrak{M}_H, w \vDash \varphi$. HL formulae are satisfied if the following holds, for any $\varphi, \psi \in$ WFF; w, w' $\in$ W; $i \in$ NOM and propositional variable p (Blackburn, 2000, 347; Areces and ten Cate, 2007, 825):

1. $\mathfrak{M}_H, w \vDash p$ iff $w \in V_H(p)$,
2. $\mathfrak{M}_H, w \vDash i$ iff $w \in V_H(i)$,
3. $\mathfrak{M}_H, w \vDash \neg\varphi$ iff $\mathfrak{M}_H, w \nvDash \varphi$,
4. $\mathfrak{M}_H, w \vDash \varphi \wedge \psi$ iff $\mathfrak{M}_H, w \vDash \varphi$ and $\mathfrak{M}_H, w \vDash \psi$,
5. $\mathfrak{M}_H, w \vDash \varphi \rightarrow \psi$ iff $\mathfrak{M}_H, w \nvDash \varphi$ or $\mathfrak{M}_H, w \vDash \psi$,
6. $\mathfrak{M}_H, w \vDash \Box\varphi$ iff $\forall w' \in W$ if wRw' then $\mathfrak{M}_H, w' \vDash \varphi$,
7. $\mathfrak{M}_H, w \vDash \Diamond\varphi$ iff $\exists w' \in W$ such that wRw' and $\mathfrak{M}_H, w' \vDash \varphi$,
8. $\mathfrak{M}_H, w \vDash @_i\varphi$ iff $\mathfrak{M}_H, w' \vDash \varphi$, where $w' \in V_H(i)$.

Conditions 1., 3., 4., 5., 6., and 7. are akin to K and K_T ones. The only two different are 2. and 8. What 2. states is a nominal i is satisfied at world w if and only if i is the name of w, that is, if w and i are equal. 8., on its part, states that $@_i\varphi$-like formulae are satisfied at some world if and only if φ is satisfied at world denoted by the nominal.

If φ is satisfied at every possible worlds of every model $\mathfrak{M}_H$ based on $\mathfrak{F}$ then we say φ is valid in $\mathfrak{F}$. Logically speaking: $\mathfrak{F} \Vdash \varphi$. If φ is valid in every frame $\mathfrak{F}$ we say it is just valid, i.e., $\Vdash \varphi$.

Basic Hybrid Logic system basically consists in this. It possesses three main characteristics: firstly, as well as in Modal Logic, it is decidable in PSPACE; secondly, it can be translated to First Order Logic if we extend the Standard Translation to nominals and satisfaction operators; and thirdly, it constitutes the FOL invariant for bisimulation fragment of formulae with constants and identity relation.

Let us focus on the second and the third characteristics. In order to express by means of First Order Correspondence Language Hybrid Logic formulae we just have to extend the Standard Translation conditions (see page 45) to represent nominals and satisfaction operators. These two conditions may be (Blackburn 2006, 344):

7. $\mathrm{ST}_x(i) = (x = i)$,
8. $\mathrm{ST}_x(@_i\varphi) = \mathrm{ST}_i(\varphi)$.

In FOCL nominal i is symbolized by constant i whereas the satisfaction operator is represented by substituting the free variable x by the constant i.

(4*) Standard Translation would then be:

$$\Diamond(i \wedge \mathrm{p}) \wedge \Diamond(i \wedge \mathrm{q}) \rightarrow \Diamond(\mathrm{p} \wedge \mathrm{q}) =$$
$$\mathrm{ST}_x(\Diamond(i \wedge \mathrm{p}) \wedge \Diamond(i \wedge \mathrm{q}) \rightarrow \Diamond(\mathrm{p} \wedge \mathrm{q})) =$$
$$\exists y\,(\mathrm{R}^{\mathrm{m}}\langle x, y\rangle \wedge (y = i \wedge \mathrm{P}(y))) \wedge \exists y\,(\mathrm{R}^{\mathrm{m}}\langle x, y\rangle \wedge (y = i \wedge \mathrm{Q}(y))) \rightarrow$$
$$\exists y\,(\mathrm{R}^{\mathrm{m}}\langle x, y\rangle \wedge (\mathrm{P}(y) \wedge \mathrm{Q}(y))).$$

As in ML and TL, Standard Translation also preserves satisfiability so that:

> **Definition 4.2.1. HL-FOCL Equisatisfiablity (HL ≈ FOCL)** For any basic hybrid formula φ, any model $\mathfrak{M}_{\mathrm{H}}$ and any possible world w in $\mathfrak{M}_H$, $\mathfrak{M}_H$, $\mathrm{w} \vDash \varphi$ if and only if $\mathfrak{M}_H \vDash \mathrm{ST}_x(\varphi)\ [x \leftarrow \mathrm{w}]$.

Therefore, from **HL ≈ FOCL** is deduced that

$$\Diamond(i \wedge \mathrm{p}) \wedge \Diamond(i \wedge \mathrm{q}) \rightarrow \Diamond(\mathrm{p} \wedge \mathrm{q})$$

and

$$\mathrm{ST}_x(\Diamond(i \wedge \mathrm{p}) \wedge \Diamond(i \wedge \mathrm{q}) \rightarrow \Diamond(\mathrm{p} \wedge \mathrm{q}))$$

are equisatisfiables.

If so, then HL is a FOL fragment, as we have declared. In particular, it is the fragment of any invariant for bisimulation formula composed by constants and identity relation. The following proposition reflects that:

Definition 4.2.2. HL Bisimulation If $\mathfrak{M}_H$ and $\mathfrak{M}_H$* are two basic Hyrbid Logic models such that $\mathfrak{M}_H = \langle W, R, V_H \rangle$ and $\mathfrak{M}_H{}^* = \langle W', R', V'_H \rangle$ then the relation $B \subseteq W \times W'$ is a bisimulation between $\mathfrak{M}_H$ and $\mathfrak{M}_H$* if the following conditions are satisfied:

1. If wBw' then, for any $p \in PROP$, $w \in V_H(p)$ if and only if $w' \in V'_H(p)$.
2. If wBw' and wRv then w'R'v' and vBv'.
3. If wBw' and w'R'v' then wRv and vBv'.
4. If, for any $w \in W$ and $w' \in W'$, w and w' are denoted by the same nominal then wBw'.

It is said a FOL formula φ with one free variable is invariant for bisimulation if its valuation at bisimilar worlds is always the same. That HL is a proper fragment of FOL is derived from this definition and from the *Characterization Theorem* (Blackburn 2006, 345; Areces and ten Cate 2007, 838-839), stating that if a first order formula φ with one free variable is invariant for bisimulation then φ is equivalent to its corresponding standard translation formula in HL.

The above system is basic for it just adds NOM and @ to ML language. However, Prior´s one is stronger. Despite us, who have drawn from Modal Logic to build Hybrid Logic, Prior draws from Temporal Logic and extends it by adding three elements: nominals, the Universal Modality and the ∀ and ∃ quantifiers.

Regarding nominals, it cannot be said any more but he adds a new operator, Q, as an alternative to them. Hence, a formula as $Q\varphi$ is true at any point of the model if and only if there is just one point on it where φ is true. That means that, through $Q\varphi$, φ is transformed into a nominal, for this is equivalent to say that φ is true at w if and only if w belongs to the set of points satisfying φ. As there is just one, φ and w are identical.

The Universal Modality is a tool which lies in most of the contemporary Hybrid Logic systems and which possesses two forms: one equivalent to □, symbolized as A, and another equivalent to ◇, symbolized as E. Aφ means that φ is true at every world of the model, whereas Eφ means that φ is true at some world in the model.

As it is easily seen, $E(i \wedge \varphi)$ truth condition is equivalent to $@_i\varphi$ one and thus $@_i\varphi$ may be defined in terms of E as:

$$@_i\varphi \stackrel{\text{def}}{=} E(i \wedge \varphi),$$

which states there exists some point in the model where *i* and φ are simultaneously true. But $@_i\varphi$ may be also defined in terms of A as follows:

$$@_i\varphi \stackrel{\text{def}}{=} A(i \rightarrow \varphi),$$

which sets that at every world of the model where i is true φ is true too.

Finally, Prior adds the quantifiers ∀ and ∃, notwithstanding the novelty that, besides their regular use, they can bind nominals. In this way, formulae such as

$$\exists x@_x\varphi$$

or

$$\forall x@_x\mathrm{F}\varphi$$

could also be included in the strong Hybrid Logic language (HL^+), named $\mathcal{L}_{HL+}$.

$\mathcal{L}_{HL+}$ is composed, apart for PROP, NOM and MOD, by a new sort of state variables SVAR = $\{x, y, z, ...\}$ representing nominals. The difference between NOM and SVAR lies in NOM elements are constants, they always denote the same world, whereas SVAR ones are variables. Hence, $\mathcal{L}_{HL+}$ is defined as (Blackburn 2006, 351):

$$x \mid i \mid \mathrm{p} \mid \neg\varphi \mid \varphi \wedge \psi \mid \varphi \rightarrow \psi \mid \Box\varphi \mid \Diamond\varphi \mid @_i\varphi \mid \mathrm{A}\varphi \mid \mathrm{E}\varphi \mid \forall x\varphi \mid \exists x\varphi.$$

Many of these operators may be interdefined (the rest of logical constants and $\Box\varphi \equiv \neg\Diamond\neg\varphi$, $@_i\varphi \equiv \mathrm{E}(i \wedge \varphi)$, $\mathrm{A}\varphi \equiv \neg\mathrm{E}\neg\varphi$, $\forall x\varphi \equiv \neg\exists x\neg\varphi$, and vice versa), but we have chosen for adding them all in order to fully present HL^+ alphabet.

With respect to its semantics, frames and models still are Kripkean, i.e., $\mathfrak{M}_{H+} = \langle \mathrm{W}, \mathrm{R}, \mathrm{V}_{H+} \rangle$. However, as now there are free and bound variables it is necessary to assign truth values in accordance with SVAR and consequently to relativize formulae valuation to variables assignments.

Let g be a function assigning values in $\mathfrak{M}_{H+}$ to variables. g is a function from SVAR to W, that is, g: SVAR → W and moreover, if g and g' are functions assigning values to variables in $\mathfrak{M}_{H+}$ but g possibly differs from g' in x value, we say that g' is a x-variant of g, and we denote it as $g' \sim^x g$. We also represent that any formula φ is satisfied at world w of model $\mathfrak{M}_{H+}$ under the assignment g as $\mathfrak{M}_{H+}, g, \mathrm{w} \vDash \varphi$.

Satisfiability conditions are defined as follows:

1. $\mathfrak{M}_{H+}, g, \mathrm{w} \vDash x$ iff $\mathrm{w} = g\,(x)$,
2. $\mathfrak{M}_{H+}, g, \mathrm{w} \vDash i$ iff $\mathrm{w} \in \mathrm{V}_{H+}(i)$,
3. $\mathfrak{M}_{H+}, g, \mathrm{w} \vDash \mathrm{p}$ iff $\mathrm{w} \in \mathrm{V}_{H+}(\mathrm{p})$,
4. $\mathfrak{M}_{H+}, g, \mathrm{w} \vDash \neg\varphi$ iff $\mathfrak{M}_{H+}, g, \mathrm{w} \nvDash \varphi$,
5. $\mathfrak{M}_{H+}, g, \mathrm{w} \vDash \varphi \wedge \psi$ iff $\mathfrak{M}_{H+}, g, \mathrm{w} \vDash \varphi$ and $\mathfrak{M}_{H+}, g, \mathrm{w} \vDash \psi$,
6. $\mathfrak{M}_{H+}, g, \mathrm{w} \vDash \varphi \rightarrow \psi$ iff $\mathfrak{M}_{H+}, g, \mathrm{w} \nvDash \varphi$ or $\mathfrak{M}_{H+}, g, \mathrm{w} \vDash \psi$,
7. $\mathfrak{M}_{H+}, g, \mathrm{w} \vDash \Box\varphi$ iff $\forall \mathrm{w}' \in \mathrm{W}$ if wRw' then $\mathfrak{M}_{H+}, g, \mathrm{w}' \vDash \varphi$,

8. $\mathfrak{M}_{H+}, g, \mathrm{w} \vDash \Diamond\varphi$ iff $\exists \mathrm{w}' \in \mathrm{W}$ such that wRw' and $\mathfrak{M}_{H+}, g, \mathrm{w}' \vDash \varphi$,
9. $\mathfrak{M}_{H+}, g, \mathrm{w} \vDash @_i\varphi$ iff $\mathfrak{M}_{H+}, g, \mathrm{w}' \vDash \varphi$, where $\mathrm{w}' \in \mathrm{V}_H(i)$,
10. $\mathfrak{M}_{H+}, g, \mathrm{w} \vDash \mathrm{A}\varphi$ iff $\forall \mathrm{w}' \in \mathrm{W}\ \mathfrak{M}_{H+}, g, \mathrm{w}' \vDash \varphi$,
11. $\mathfrak{M}_{H+}, g, \mathrm{w} \vDash \mathrm{E}\varphi$ iff $\exists \mathrm{w}' \in \mathrm{W}\ \mathfrak{M}_{H+}, g, \mathrm{w}' \vDash \varphi$,
12. $\mathfrak{M}_{H+}, g, \mathrm{w} \vDash \forall x\varphi$ iff, for any $g' \sim^{\mathrm{x}} g$, $\mathfrak{M}_{H+}, g', \mathrm{w} \vDash \varphi$,
13. $\mathfrak{M}_{H+}, g, \mathrm{w} \vDash \exists x\varphi$ iff, for some $g' \sim^{\mathrm{x}} g$, $\mathfrak{M}_{H+}, g', \mathrm{w} \vDash \varphi$.

2.-9. conditions are similar to the ones we have already seen. The new ones are 1., 10., 11., 12. and 13. 1. sets that if variable x is true at w according to g it is because the value assigned to x by g is w. On their part, 10.-13. are no more than the formal representation of what we have pointed out above.

$\mathcal{L}_{HL+}$ and $\mathfrak{M}_{H+}$ may be extended to Temporal Logic and to accept propositions composed by temporal operators whose valuation is carried out at instants and not at worlds. And $\mathcal{L}_{HL+}$ can also be extended by means of other operators such as ↓, **Until** and **Since** (Areces and ten Cate 2007, 822-823) to increase its expressive power.

Anyway, the most relevant thing about strong Hybrid Logic is that, apart from it is still decidable despite its huge expressive capacity, it is also powerful enough to translate FOCL to its own language. Until now we have claimed that FOCL is able to translate ML, TL and basic HL formulae by means of Standard Translation but HL^+ is, in turn, able to translate FOCL formulae thanks to what it is called *Hybrid Translation*.

Let HT be a function assigning to each FOCL formula its corresponding HL^+ one. HT consists in the following, for any variable $x, y, z \in$ SVAR, any propositional symbol $\mathrm{p} \in$ PROP and any $\varphi, \psi \in$ WFF (Blackburn 2006, 352):

1. $\mathrm{HT}(\mathrm{R}^m\langle x, y\rangle) = @_x\langle \mathrm{m}\rangle y$,
2. $\mathrm{HT}(\mathrm{P}(x)) = @_x\mathrm{p}$,
3. $\mathrm{HT}(x = y) = @_x y$,
4. $\mathrm{HT}(\neg\varphi) = \neg\mathrm{HT}(\varphi)$,
5. $\mathrm{HT}(\varphi \wedge \psi) = \mathrm{HT}(\varphi) \wedge \mathrm{HT}(\psi)$,
6. $\mathrm{HT}(\varphi \rightarrow \psi) = \mathrm{HT}(\varphi) \rightarrow \mathrm{HT}(\psi)$,
7. $\mathrm{HT}(\exists z\varphi) = \exists z\mathrm{HT}(\varphi)$,
8. $\mathrm{HT}(\forall z\varphi) = \forall z\mathrm{HT}(\varphi)$.

Similarly for the rest of formulae.

As it may be observed the main key of Hybrid Translation is @. Thanks to it we can reduce FOCL to HL^+, so without @ Hybrid Logic would not be capable enough of carrying out this translation.

Hybrid Logic (both basic and strong) basically consists in this, and its main advantage over the rest logics we have presented here is Hybrid Translation.

5. Conclusion

As we asserted at the Introduction this paper aims to show how Hybrid Logic was developed by Arthur Prior in order to be an extent of Temporal Logic to solve its problems (and thus Modal Logic ones) and why it is stronger than both of them.

With the aim of fulfilling this goal in section 2 we exposed ML and TL to see, in the first place, how Temporal Logic is built by Prior from Modal Logic. Compared with ML, TL constitutes a huge advance at formally representing propositions suffused with temporal information, but it suffers from the same Modal Logic disadvantages.

In section 3 we entered such disadvantages. They are mainly two: on the one hand, both ML and TL may be subsumed under the First Order Correspondence Language. In the first subsection, we saw that by means of the Standard Translation FOCL may translate any ML/TL formula into its own language. We explained why: for Modal Logic is the FOCL fragment composed by every invariant for bisimulation formulae. On the other hand, ML and TL are incapable of naming specific points inside a model, that is, they have no mechanism allowing to point that such and such propositions is true at some particular point. First Order Logic does have, however.

That means FOL is stronger than ML and TL. Nevertheless, ML and TL main advantage is they provide an internal perspective of models in contrast to FOL, whose perspective is external. That is precisely why at the end of the second subsection of section 3 we claimed that both are better to propose a logical system which does its best to reflect the natural concept of time. It is just necessary to combine ML/TL mechanisms with FOL ones to achieve it. The result of such combination is Hybrid Logic.

Section 4 has been devoted entirely to it. In the first part, we explained its motivation. The need to pose a more powerful system than Temporal Logic lied in the Priorean approach regarding Time B series may be reduced to A ones. As TL language is not expressive enough to allow such thing, by means of the four grades of tense-logical involvement we saw how Prior formulates a system which does allow to reduce First Order Logic to a purely temporal logic. A reduction which can only be carried out by HL mechanisms.

In the second part of this section we exposed such mechanisms. The use of a new set of symbols representing nominals and of satisfaction operators constitute the Hybrid Logic main achievements. Thanks to them, and to operators such as A, E, ∀ and ∃, we are able to develop an even stronger system than HL to which it may now be translated the First

Order Correspondence Language. At the end of this part it has been indicated how to do that by using Hybrid Translation.

Hybrid Translation is the most relevant result proposed in this paper. By means of it, it is proved that HL is an incredibly powerful system due to its huge expressivity. An expressiveness which, as pointed, allows to translate FOCL to HL language. And since Standard Translation allows to do the same with ML/TL languages then both logics may also be translated to Hybrid one.

In formal terms, we could say:

$$ML/TL \subseteq FOCL \subseteq HL,$$

that is, Modal Logic and Temporal Logic are properly included in the First Order Correspondence Language, which in turn it is included (but not properly) in Hybrid Logic. So, both FOCL and HL may translate each other; a thing that ML and TL are not able to do with respect to FOCL. Hence Hybrid Translation relevance.

Throughout the entire exposition we have always delivered examples with the aim of clarifying the questions at stake. Therefore, it only remains to conclude by pointing out that Hybrid Logic may be applied not only to what we have mentioned but "hybridization" may be applied to a great range of contexts.

As we said in section 4.1., Prior realized that HL mechanisms allow to consider as propositions not only predicates alluding to instants but just any predicate. In this paper, we have exclusively focused on Kripke semantics, but Hybrid Logic can also be extended to topological and algebraic semantics. HL tools may be added both to basic Modal Logic and to First Order and Second Order Modal Logics (intuitionism-based even), allowing to obtain very interesting results. An example may be found on Carlos Areces, Patrick Blackburn, Antonia Huertas and María Manzano´s (2011) paper where a Hybrid Types Theory is developed. Or on Areces, Blackburn, Huertas and Manzano´s (2014) paper where a Completeness proof, via Henkin, of this Hybrid Types Theory is provided.

Some of these applications have already been investigated, but there are many others which remain unstudied. In our case, we think it would be so interesting due to its pragmatic involvements to combine Hybrid Logic with Epistemic Logic and Fuzzy one in order to build models reflecting more accurately the way we, human beings, reason and argue.

When we utter «I think φ» or «I know φ» this believe or knowledge is usually context-dependent, i.e., we can believe or know φ in some particular moment, but not to believe it or know it in another one (for our set of knowledges has changed, for instance). And even more. In may times we do not totally believe or know φ. We may have doubts about its truth (or falsehood) and not to believe it *too much*, or not to know it *very well*.

Classical Epistemic Logic allows to formalize propositions such as «I think φ» or «I know φ» and to valuate them according to the epistemic states of some model. By combining it with Fuzzy Logic we obtain a many-valued system allowing to increase the number of truth values we may assign to formulae to represent values such as *too much* or *not too much*, *very well* or not *very well*, *a little*, *enough*, etc. And if we add all this to Hybrid Logic we could build a system in which, moreover, it could be possible to reflect the fact that certain propositions possess just a truth value at some specific point. For instance, we could formally reflect and valuate a statement as (9) «I am not quite sure in this moment of φ».

Utterances as (9) are quite common in our daily communicative exchange. Our believes and knowledges usually tend to be fuzzy and context-dependents, and that is why we ought to appeal to mechanisms from these three systems (epistemic, fuzzy and hybrid) if we wish to develop models which depict such characteristics. Models whose application may go from Knowledge and Reasoning Representation in Artificial Intelligence to Argumentation.

6. Appendix

6.1. Notation

The syntax of the logical language $\mathcal{L}$ used in this paper is composed by:

A. Primitive Symbols

1. Propositional variables: p, q, r, ... That is, lowercase letters of the alphabet beginning from p.
2. Connectives:
    i. Negation: $\neg$,
    ii. Conjunction: $\wedge$,
    iii. Disyunction: $\vee$,
    iv. Material conditional: $\rightarrow$,
    v. Logical equivalence: $\equiv$.
3. Quantifiers:
    i. Universal: $\forall$,
    ii. Existential: $\exists$.

B. Metavariables

Representing either propositional variables or well-formed formulae: φ, ψ, ...

C. Well-Formed Formulae (WFF)

Which are any formulae fulfilling the following conditions:

1. Both propositional variables and metavariables are wff.
2. If p is a wff, then $\neg$p is a wff too.
3. If φ and ψ are wff, then $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, $(\varphi \rightarrow \psi)$, $(\varphi \equiv \psi)$, $\forall\varphi$ and $\exists\varphi$ are wff too.
4. The outermost parentheses of any well-formed formula may be omitted.
5. Anything not followed from the recursive application of rules 1.-4. is a wff.

We shall call PROP to the set of propositional variables and WFF to the set of well-formed formulae. Besides, apart from these symbols over the course of our paper we have introduced new ones inasmuch as we have needed it.

6.2. Abbreviation Index

We have used the following abbreviations throughout the paper:

- FOCL for *First Order Correspondence Language*.
- FOL for *First Order Logic*.
- HL for *Hybrid Logic*.
- HL^+ for *Strong Hybrid Logic*.
- ML for *Modal Logic*.
- TL for *Temporal Logic*.

References

Areces, C. and ten Cate, B. (2007). Hybrid Logics. In P. Blackburn, J. van Benthem and F. Wolter (eds.). *Handbook of Modal Logic*, Vol. 3, pp. 822-863. Amsterdam: Elsevier.

Areces, C., Blackburn, P., Huertas, A. and Manzano, M. (2014). Completeness in Hybrid Type Theory. *Journal of Philosophical Logic*, 43(2-3): 209-238.

Areces, C., Blackburn, P., Huertas, A. and Manzano, M. (2011). Hybrid Type Theory: A Quartet in Four Movements. *Principia: An International Journal of Epistemology*, 15(2): 225-247.

Blackburn, P. (2000). Representation, Reasoning, and Relational Structures: a Hybrid Logic Manifesto. *Logic Journal of the IGPL*, 8(3): 339-365.

Blackburn, P. (2006). Arthur Prior and Hybrid Logic. *Synthese*, 150(3): 329-372.

Blackburn, P. and van Benthem, J. (2007). Modal Logic: A Semantic Perspective. In P. Blackburn, J. van Benthem and F. Wolter (eds.), *Handbook of Modal Logic*, Vol. 3, pp. 2-79. Amsterdam: Elsevier.

Findlay, J. (1941). Time: A Treatment of Some Puzzles. *Australasian Journal of Psychology and Philosophy*, 19: 216-235.

McTaggart, J. (1908). The Unreality of Time. *Mind*, 17(68): 457-474.

Øhrstrøm, P. and Hasle, p. (1995). *Temporal Logic. From Ancient Ideas to Artificial Intelligence*. Dordrecht: Springer Science+Business Media Dordrecht.

Prior, A. (1955). Diodoran Modalities. *The Philosophical Quarterly*, 5(20): 205-213.

Prior, A. (1957). *Time and Modality*. Oxford: Oxford University Press.

Prior, A. (1967). *Past, Present and Future*. Oxford: Oxford University Press.

Prior, A. (2010). *Papers on Time and Tense*. Oxford: Oxford University Press.

Reichenbach, H. (1947). *Elements of Symbolic Logic*. New York: Dover Publications.

Rescher, N. and Alasdair, U. (1971). *Temporal Logic*. New York: Springer.

van Eijck, J. (2006). Bisimulation. In https://homepages.cwi.nl/~jve/courses/lai0506/LAI11.pdf.

Revista de Humanidades de Valparaíso
Año 7, 2019, 1er semestre, No 13, págs. 68-89
eISSN 0719-4242

Humanities Journal of Valparaiso
No 13 (2019): 68-89

Giro dinámico y lógica de la investigación científica[†]

Dynamic turn and logic of scientific research

Mª Victoria Murillo-Corchado[*]; Ángel Nepomuceno-Fernández[**]

[*]Universidad de Sevilla, España
mvmurilloc@gmail.com

[**]Universidad de Sevilla, España
nepomuce@us.es

DOI: https://doi.org/10.22370/rhv2019iss13pp68-89

Recibido: 25/07/2019. Aceptado: 16/08/2019

Resumen

Para presentar la incidencia del giro dinámico en la lógica de la investigación científica, en este artículo comenzamos con una sección que trata de los juegos lógicos como desencadenantes de este giro dinámico en la lógica contemporánea, junto con el programa de dinámica lógica de la información y la interacción. Sucintamente presentamos las principales características de la lógica favorable a la independencia y la semántica juego-teórica (IF-logic y GTS, respectivamente, en Hintikka y Sandu 1997), de la lógica dialógica (Redmond y Fontaine 2011), así como los elementos esenciales de dicho programa. Si bien a partir de cualquiera de estos puntos de vista se cuenta con un elenco de herramientas lógicas para abordar cuestiones más claramente epistemológicas, destacamos el papel de la lógica epistémica dinámica (LED), a la que dedicamos la siguiente sección. Sigue otra en la que entramos en los estudios lógicos de la abducción como uno de los problemas fundamentales de la epistemología contemporánea y, en una nueva sección, en términos de los constructos teóricos de las secciones precedentes, presentamos y explicamos un fenómeno surgido en el campo de la lingüística, el caso del descubrimiento de la lengua amazónica pirahã, que debería ser considerado una anomalía en el marco de la teoría chomskiana.

[†]Este trabajo se ha realizado en el marco de las actividades que lleva a cabo el Grupo de Investigación en Lógica, Lenguaje e Información de la Universidad de Sevilla (GILLIUS http://grupo.us.es/ghum609/php/node/41), referencia HUM-609, del Plan Andaluz de Investigación, Desarrollo e Innovación.

Palabras clave: juegos lógicos, lógica dialógica, LED, abducción, gramática universal, pirahã, anomalía, novedad.

Abstract

In order to present the incidence of the dynamic turn in the logic of scientific research, we begin with a section, in this article, that deals with logical games as triggers of this dynamic turn in contemporary logic, together with the program of logical dynamics of information and interaction. We briefly introduce the main characteristics of the logic favorable to independence and the game-theoretical semantics (IF-logic and GTS, respectively, in Hintikka and Sandu 1997), of dialogical logic (Redmond and Fontaine 2011), as well as the essential elements of this program. Although from any of these points of view we have a list of logical tools to deal with more clearly epistemological issues, we highlight the role of dynamic epistemic logic (DEL), to which we dedicate the following section. There follows another one in which we enter the logical studies of abduction as one of the fundamental problems of contemporary epistemology and, in a new section, in terms of the theoretical constructs of the preceding sections, we present and explain a phenomenon that emerged in the field of linguistics, the case of the discovery of the Amazonian language pirahã, which should be considered an anomaly within the framework of the Chomskian theory.

Keywords: Logical games, dialogical logic, DEL, abduction, universal grammar, pirahã, anomaly, novelty.

1. Introducción

El *giro dinámico* en lógica, por cuanto éste representa en términos kuhnianos un cambio de paradigma en la investigación lógica actual, tiene lugar a partir de la aparición de la semántica juego-teórica (GTS) según señala P. Gochet: "The invention of game-theoretic semantics by J. Hintikka in the seventies (Hintikka 1973; Saarinen 1979) can be described as the emergence of a new paradigm not only for semantics but also for logic and the philosophy of mathematics" (Gochet 2002, 175). Aceptando esta observación, añadimos que el giro en cuestión tiene como ingredientes, de un lado, las propuestas de estudiar las tareas lógicas como si éstas fueran desarrollos de formas de juego y, de otro, el vasto programa denominado "dinámica lógica de la información y de la interacción", propuesto en van Benthem (2011).

Por otra parte, las bases para una lógica de la investigación científica ya no se establecen, bajo el supuesto de que ello es posible, exclusivamente en términos de lógica clásica. Más bien son las lógicas no clásicas las que proporcionan las herramientas necesarias para el abordaje de este tipo de problemas epistemológicos. La noción de abducción debida a Peirce, a pesar de su carácter poliédrico, permite avanzar en las relaciones entre

lógica y epistemología, de manera que, si hacemos la tradicional distinción entre contexto de descubrimiento y contexto de justificación, el primero no es ajeno a la lógica misma, en contra de la opinión de autores como Popper, según el cual el contexto de descubrimiento no requería un análisis lógico por ser temática más bien de la psicología experimental (Popper 1980: 30-31). Ahora bien, "en la epistemología de Peirce, el pensamiento es un proceso dinámico, esencialmente una acción que oscila entre los estados mentales de duda y de creencia" (Aliseda 2014, 46).

2. Juegos lógicos

Al producirse el giro dinámico, surgen varias propuestas para entender las tareas lógicas en forma de juegos, con lo que las nociones lógicas básicas emergen uniformemente como estrategias ganadoras en diversos escenarios. Las principales propuestas son:

1. Hintikka, GTS. Se trata de hacer una evaluación de juegos entre un *Verificador* y un *Falsificador*, teniendo en cuenta un oráculo (un modelo).
2. Ehrenfeucht-Fraïssé. Se trata de juegos entre un *Duplicador* y un *Estropeador* —*spoiler*— que comparan modelos.
3. Lorenz-Lorenzen. Establecen juegos de diálogos argumentativos entre un *Proponente* y un *Oponente*.

Uno de los ingredientes del caldo de cultivo que propicia la aparición del nuevo paradigma es el origen mismo de los computadores modernos y el diseño de máquinas de deducción lógica. Surge así un extraordinario interés por los fenómenos de la comunicación en general, un mundo cognitivo en el que se puede situar el dominio de la lógica, como "el estudio de los invariantes subyacentes en estos procesos informacionales" (van Benthem 2011, 1). Así pues, son una gama de procesos informacionales los que deben constituir la teoría lógica. Ahora bien, para trabajar en este ámbito se plantea la necesidad de abordar la interacción entre estructuras estáticas y dinámicas.

Se puede decir que tanto la lógica clásica como los avances en la tradición algebraica convergen finalmente en los planteamientos de los juegos lógicos y la dinámica lógica de la información, aunque se ha señalado a Peirce como un precedente inmediato. En efecto, para el lógico norteamericano el hablante es esencialmente un defensor de su propia proposición y su deseo es interpretarla de manera que sea defendible, mientras que el intérprete no está tan interesado en interpretarla plenamente sin considerar a qué extremo se puede llegar, es relativamente hostil y busca la interpretación menos defendible.

Considerada la actividad lógica como un juego —en este punto Hintikka se considera un seguidor de Wittgenstein, posición discutible para algunos (Marion 2006)—, se había de replantear el sentido de las reglas de inferencia. Para justificar una inferencia, Hintikka propone dos clases de reglas o principios, a saber, las reglas definitorias, similares a las que definen un juego, como, por ejemplo, en el ajedrez —la deducción o la indagación

científica contienen los elementos de un juego estratégico—. Estas reglas nos indican los movimientos posibles en una situación dada (en el curso de un juego como el ajedrez, este tipo de reglas nos dicen cómo se mueven la reina, el alfil, la torre, etc.). Hay otras reglas, las estratégicas, que nos indican cuáles son los mejores movimientos para ganar el juego (o evitar que nos gane el contrincante; en el caso del ajedrez, cómo evitar un jaque pastor, etc.). Además, Hintikka establece una lógica de cuestiones y respuestas, que tiene en cuenta las diferencias entre razonamiento ampliativo y no ampliativo —distinción que se da entre pasos de un argumento: interrogativo (ampliativos) y deductivos (no ampliativos)—. En la indagación interrogativa la cuestión es anticipar la situación epistémica provocada por la respuesta. Todas estas observaciones deberían tenerse en cuenta como un importante conjunto de consejos certeros para abordar aproximaciones lógicas a la inferencia científica.

Desde la GTS propuesta para interpretar la lógica favorable a la independencia (If-logic, Hintikka 1997; Hintikka y Sandu 1997; Aho y Pietarinen 2006), donde el lenguaje se considera orientado a metas más que como una actividad gobernada por reglas, conocer el significado de una oración es conocer el cambio en el estado de información, que parece ampliar. A partir de la idea de actualización de la información, cabe definir una noción de actualización mínima (van Benthem 2011), siempre que los estados cognitivos (o estados de información) sean ordenados por inclusión. A diferencia de la semántica tarskiana, que adopta el *principio de composicionalidad* —el significado de una proposición compleja viene dado por el significado de las proposiciones más simples que la integran—, el punto de vista de la GTS toma como base el *principio de significado dependiente del contexto* y considera que éste resulta incompatible con el principio de composicionalidad —el significado de una proposición compleja es una función del significado de sus partes constituyentes—, incompatibilidad que podría ser eliminada cambiando la concepción estática del significado por otra de carácter dinámico.

Dados un modelo M y una fórmula φ (para simplificar, nos mantendremos a nivel proposicional), un *juego semántico asociado* al modelo M y la fórmula φ, en símbolos G (M, φ), caracteriza la verdad o falsedad de la fórmula en el juego semántico, que es jugado por dos jugadores, *Abelardo* y *Eloísa*, que se pueden representar como A y E, respectivamente. En el juego G (M, φ) tanto A como E han de hacer una elección, de acuerdo con unas reglas establecidas para cada clase de fórmulas, es decir por las reglas constitutivas de G (M, φ). Estas son:

1. φ es atómica. No se hace ningún movimiento:
    a. Si $M \vDash \varphi$, entonces E ha ganado,
    b. Si $M \nvDash \varphi$, entonces A ha ganado.
2. $\varphi = \delta \vee \psi$: E elige $\theta \in \{\delta, \psi\}$ y el juego continúa como G (M, θ).
3. $\varphi = \delta \wedge \psi$: como antes, sólo que A es quien hace la elección.

4. $\varphi = \neg\psi$: lo mismo que en el juego G (M, ψ), excepto que los papeles de los jugadores A y E cambian, incluyendo las reglas de ganancia y pérdida.

Asimismo, hemos de considerar las reglas de estrategia, teniendo en cuenta que cada juego G (M, φ) es un juego de *información perfecta,* los jugadores "conocen" todas las elecciones hechas en el desarrollo del juego —se pueden considerar también de información imperfecta—. Al final del juego, uno de los jugadores gana y el otro pierde, de acuerdo con la circunstancia de que la fórmula atómica con la cual termina el juego sea verdadera o falsa en el modelo (que pertenezca o no a M). Una estrategia de un jugador X en un juego G (M, φ) no es más que un método que produce un movimiento (legal) para X contra cualquier secuencia de movimientos hecha por su oponente. Para una fórmula φ y un modelo M, una estrategia para un jugador X es un conjunto F_i de funciones f_Q, correspondientes a las distintas constantes lógicas Q, que pueden incitar a un movimiento de X en G (M, φ). Es una estrategia ganadora para el jugador X, si lleva a la victoria de X contra cualquier movimiento del oponente.

El estudio del valor de verdad se hace atendiendo a las siguientes estipulaciones, para un modelo M y una fórmula φ:

1. φ es verdadera (en el sentido de la semántica juego-teórica —GTS—) syss existe una estrategia ganadora para E en el juego G (M, φ); simbólicamente:

$$M \models^{+}_{GTS} \varphi$$

2. φ es falsa (en el sentido de la semántica GTS) syss existe una estrategia ganadora para A en G (M, φ); simbólicamente:

$$M \models^{-}_{GTS} \varphi$$

El otro punto de vista significativo es el de la elaboración de una lógica dialógica. Se puede decir que existe una relación entre los diálogos y las reglas de razonamiento correcto (dialéctica, *obligaciones*, etc.). A mediados del s. XX, P. Lorenzen puso en relación diálogos y fundamentos constructivos de la lógica. Se define un diálogo D sobre una proposición (en su caso, una fórmula) φ —D (φ)—, que comienza con φ afirmada por un jugador, y alcanza una posición final con victoria o derrota después de un número finito de movimientos de acuerdo con reglas definidas (Rahman, Clerbout y Keiff 2009; Redmond y Fontaine 2011). La noción dialógica de demostración descansa en esta noción de juego. En cuanto a la estructura de un juego dialógico, se considera que en el diálogo participan un proponente (P) y un oponente (O), que discuten una tesis siguiendo ciertas reglas, ejecutando los movimientos de ataque o de defensa. Las reglas se dividen en:

1. De partícula (de las constantes lógicas), que muestran qué movimientos están autorizados para atacar los movimientos del otro jugador o defender los propios.
2. Estructurales, que determinan el curso general del juego dialógico. Estas a su vez son:

a. Inicial: la fórmula es afirmada por P. Alternativamente habrá movimientos de P y de O. Cada movimiento que sigue al inicial es un ataque o una defensa.

b. Situación: P y O sólo pueden hacer movimientos que cambien la situación.

c. Formal: P no puede introducir fórmulas atómicas, salvo que previamente las haya afirmado O.

d. Ganancia: X gana syss es el turno de Y pero éste no puede hacer ningún movimiento.

e. Intuicionista: en cada movimiento, cada jugador puede atacar una fórmula compleja afirmada por el otro o puede defenderse contra el último ataque que no ha sido aún defendido.

f. Clásica: en cualquier movimiento un jugador puede atacar una fórmula compleja afirmada por el otro o puede defenderse contra cualquier ataque (incluyendo aquellos que ya han sido defendidos).

En la construcción de diálogos se usará la regla intuicionista o la regla clásica, según la finalidad que se persiga. Este planteamiento no ha estado exento de críticas en las cuales se ha considerado la lógica dialógica como exótica y se ha afirmado que es una lógica constructiva. En su origen, en efecto, fue propuesta como lógica intuicionista, sin embargo los posteriores desarrollos la han convertido en una especie de marco general, a partir del cual se pueden establecer diálogos para otras lógicas (modal, epistémica, etc.) y, aunque hace uso de alguna terminología propia, no complica demasiado las prácticas más que otros procedimientos (ya sean secuentes, tableaux, etc.). Este planteamiento ofrece una dimensión pragmática de las constantes lógicas. Desde luego constituyen una proto-semántica: esquema de juego que al completarse con reglas estructurales presenta la semántica del juego. Los operadores lógicos forman juegos desde otros juegos más simples, que muestran cómo relacionar sentencias y proposiciones; la aserción de una sentencia contiene una proposición con cierta fuerza conferida por el ataque (demanda de proferir una afirmación) y la defensa (respuesta de que se puede proferir la aserción).

En cualquier caso, se trata de una importante manifestación del giro dinámico que venimos comentando y se puede considerar afín a los planteamientos de la lógica favorable a la independencia y la correspondiente GTS.

Aunque el punto de vista de los juegos lógicos ha sido fundamental para la aparición del giro dinámico en lógica, hemos de destacar el desarrollo de la lógica epistémica dinámica propiciada en el mismo como elemento esencial (Baltag y Smets 2012), sobre todo los avances logrados en el marco de la *dinámica lógica de la información* (van Benthem 2011), en el que se ha planteado un amplio proyecto de investigación a cuya base está la consideración de la inferencia como un proceso informacional. Cabe decir, a este respecto, que las actividades de inferencia, evaluación, revisión de creencias, etc.

son tan importantes como sus correspondientes productos en la configuración de la teoría lógica. Basta observar cualquier proceso informacional para comprobar que en realidad ha habido observación, comunicación e inferencia, incluso que se han barajado cuestiones y respuestas. La información inicial se va modificando con las distintas acciones. Un sencillo puzle lo ejemplifica. Sobre una reunión en la Universidad se sabe que, si asiste el decano o el vicerrector, entonces asiste el director del departamento; si no asiste el decano, entonces asiste el vicerrector; si asiste el vicerrector, no asiste el director. Usando las minúsculas iniciales correspondientes, se formalizan estas reglas como:

$$d \vee v \rightarrow t; \neg d \rightarrow v; v \rightarrow \neg t.$$

El estado inicial de información lo constituyen las 8 opciones —cada literal positivo representa que la correspondiente circunstancia se da, pero si es negativo, que no se da—:

$$\{dvt, dv\neg t, d\neg vt, \neg dvt, d\neg v\neg t, \neg dv\neg t, \neg d\neg vt, \neg d\neg v\neg t\}$$

Cada una de las afirmaciones indicadas da lugar a una actualización del estado inicial:

1. $d \vee v \rightarrow t$. Nuevo estado $\{dvt, d\neg vt, \neg dvt, \neg d\neg vt, \neg d\neg v\neg t\}$
2. $\neg d \rightarrow v$. El nuevo estado (se eliminan más opciones) es $\{dvt, d\neg vt, \neg dvt\}$
3. $v \rightarrow \neg t$. Entonces el nuevo estado es $\{d\neg vt\}$

Así pues, el resultado es que a la reunión asiste el decano, no asiste el vicerrector y también asiste el director del departamento.

Otro ejemplo habitual es el de tres individuos, sean 1, 2 y 3, integrantes de un jurado, que debe elegir entre A y B (finalistas de un concurso, o dos individuos que optan a cualquier otro tipo de premio). Cada uno apunta su voto en una nota. El secretario es distinto de 1, 2 y 3 y no vota, examina las notas y sólo dice "no hay consenso". Entonces 2 muestra a 1 su voto y declaran que ellos dos no han votado lo mismo. En ese momento, 1 y 2 desconocen el resultado. Sin embargo, 3 lo conoce perfectamente. En efecto, puesto que el voto de 3 coincide con uno de los otros dos —de acuerdo con la información hecha pública por el secretario y los otros dos miembros del jurado—, 3 ya sabe que ha ganado justo quien ha sido votado por él. Se suelen dar otros ejemplos, como el de los niños con la frente manchada, a partir de los cuales se introducen las cuestiones más relevantes que se presentan en el tratamiento lógico del conocimiento (véase Fagin et al. 1995).

En última instancia, el propio giro dinámico en lógica y semántica ha inspirado la aparición y el desarrollo de la lógica epistémica dinámica (LED), su columna vertebral, que se ha beneficiado de la confluencia de esfuerzos por parte de investigadores cuyo trabajo se realiza en diversas disciplinas, principalmente, filosofía, lingüística, álgebra, ciencias de la computación y teoría de juegos, entre otras, lo que le da un marcado sello de interdisciplinariedad.

3. Lógica epistémica dinámica

La lógica epistémica es una extensión de la lógica clásica que estudia operadores epistémicos. En esta breve presentación de LED seguimos a van Ditmarsch, van der Hoek y Kooi (2008), con ligeras modificaciones de notación. Dado un conjunto de variables proposicionales $\mathcal{P} \neq \varnothing$, un lenguaje proposicional básico para LED se define de acuerdo con la siguiente regla BNF,

$$\varphi ::= p \mid \bot \mid \neg \varphi \mid \varphi \vee \mathcal{X} \mid \varphi \wedge \mathcal{X} \mid \varphi \rightarrow \mathcal{X} \mid K_a\varphi \mid [\varphi!]\, \mathcal{X};$$

donde p es una variable proposicional; $\bot$ es una constante proposicional (contradicción); $K_a\varphi$ representa que el agente a conoce φ; $a \in G$, el conjunto de todos los agentes; $[\varphi!]\mathcal{X}$ indica que tras cada anuncio de φ, $\mathcal{X}$ es el caso. El operador de conocimiento K, como los operadores modales en general, tiene su dual, a saber $\hat{K}$, el cual, para el agente a, es definible como $\hat{K}_a \varphi = \neg K_a \neg \varphi$, que se lee hasta donde el agente a conoce, φ es posible.

La semántica para este lenguaje proposicional se suele establecer en términos kripkeanos. Un modelo

$$M = \langle W, \{\mathfrak{R}_a : a \in G\}, v\rangle,$$

Donde $W \neq \varnothing$ es el conjunto de estados (o mundos); $\{\mathfrak{R}_a : a \in G\}$ es un conjunto de relaciones definidas en W para cada individuo $a \in G$ (son las relaciones de *accesibilidad*) —escribiremos sólo $\mathfrak{R}_a$, no $\{\mathfrak{R}_a : a \in G\}$—; $v : \mathcal{P} \rightarrow 2^W$, asigna a cada variable proposicional el conjunto de estados (mundos) en los que vale la variable, de manera que $v(\bot) = \varnothing$ y $v(p) \in \wp(W)$ —o, lo que es lo mismo, $v(p) \subseteq W$—.

A veces se menciona el conjunto de estados W como el dominio del modelo, por lo que se puede decir que D (M) = W, y, puesto que M es un modelo epistémico, para referirnos al modelo epistémico en el estado w, se anotará M, w. Teniendo en cuenta estas consideraciones, a partir de un modelo epistémico $M = \langle W, \mathfrak{R}_a, v\rangle$ y el estado $s \in W$, el modelo en tal estado satisface una fórmula φ, simbólicamente, $M, s \vDash \varphi$, según el siguiente clausulado

- $M, s \vDash p$ syss $s \in v(p)$,
- $M, s \vDash \neg\varphi$ syss $M, s \nvDash \varphi$,
- $M, s \vDash \varphi \vee \psi$ syss $M, s \vDash \varphi$ o $M, s \vDash \psi$,
- $M, s \vDash \varphi \wedge \psi$ syss $M, s \vDash \varphi$ y $M, s \vDash \psi$,
- $M, s \vDash \varphi \rightarrow \psi$ syss $M, s \vDash \neg\varphi$ o $M, s \vDash \psi$,
- $M, s \vDash K_a\varphi$ syss para todo $s' \in W$, si $R_a(s, s')$, entonces $M, s' \vDash \varphi$,
- $M, s \vDash \hat{K}_a\varphi$ syss existe un $s' \in W$, tal que $R_a(s, s')$ y $M, s' \vDash \varphi$,
- $M, s \vDash [\varphi!]\mathcal{X}$ syss $M, s \vDash \varphi$ implica que $M|_\varphi, s \vDash \mathcal{X}$

teniendo en cuenta que $M|_{\varphi}$, M restringido a φ, se define $M|_{\varphi} = \langle W', \Re'_a, v' \rangle$ de manera que:

1. $W' = \{s \in W : M, s \vDash \varphi\}$,
2. $\Re'_a = \Re_a \cap (W' \times W')$, para cada $a \in G$,
3. $v'(p) = v(p) \cap W'$, para cada $p \in \mathcal{P}$.

Las relaciones de accesibilidad pueden tener determinadas características (serialidad, reflexividad, transitividad, etc.). Por otra parte, sólo se anuncian públicamente verdades, en ningún caso falsedades. De acuerdo con estas características se identifican clases de marcos de Kripke, para los cuales se establecen ciertas axiomatizaciones. Un sistema axiomático mínimo para lógica epistémica consta de los siguientes (esquemas de) axiomas y reglas (haciendo uso del lenguaje referido):

1. Todas las instancias de las tautologías proposicionales
2. $K_a(\varphi \rightarrow \psi) \rightarrow (K_a\varphi \rightarrow K_a\psi)$

Además, las reglas

- Regla de *modus ponens*: de $\varphi \rightarrow \psi$ y φ, se infiere ψ
- Regla de necesitación: de φ se infiere $K_a\varphi$

Si se añaden al sistema mínimo los axiomas que indicamos a continuación, tendremos un sistema epistémico S5 y la clase de marcos de Kripke para la semántica correspondiente, M_{S5} es aquella que contiene los marcos cuyas relaciones de accesibilidad son, además de seriales, reflexivas, transitivas y simétricas. Los nuevos axiomas, relativos al conocimiento, son:

3. $K_a\varphi \rightarrow \varphi$
4. $K_a\varphi \rightarrow K_aK_a\varphi$
5. $\neg K_a\varphi \rightarrow K_a\neg K_a\varphi$

Estos se corresponden con los habituales T, S4 y S5, respectivamente, de la lógica modal alética. El axioma 4 establece que "si el agente a conoce φ, entonces este agente conoce que conoce φ", lo que es una expresión de la *introspección positiva*, mientras que el 5, "si el agente a no conoce φ, entonces este agente sabe que no conoce φ", se refiere a la *introspección negativa*. No entramos en la discusión acerca de las razones filosóficas por las cuales se debe aceptar o rechazar la introspección, en particular la negativa; no obstante, esta consideración del conocimiento en contextos inferenciales puede resultar útil, tanto en ámbitos computacionales como epistemológicos. Nótese que estos axiomas determinan condiciones de las relaciones de accesibilidad (como reflexividad, transitividad y simetría).

El sistema axiomático indicado es ampliable para llegar a constituir un sistema básico de LED. Estos esquemas que se añaden son específicos para anuncios públicos:

6. $[\varphi!]p \leftrightarrow (\varphi \rightarrow p)$ Permanencia atómica
7. $[\varphi!]\neg\mathcal{X} \leftrightarrow (\varphi \rightarrow \neg[\varphi!]\mathcal{X})$ Anuncio y $\neg$
8. $[\varphi!](\mathcal{X} \wedge \psi) \leftrightarrow ([\varphi!]\mathcal{X} \wedge [\varphi!]\psi)$ Anuncio y $\wedge$
9. $[\varphi!]K_a\mathcal{X} \leftrightarrow (\varphi \rightarrow K_a[\varphi!]\mathcal{X})$ Anuncio y conocimiento
10. $[\varphi!][\mathcal{X}!]\psi \leftrightarrow [\varphi \wedge [\varphi!]\mathcal{X}!]\psi$ Composición de anuncios.

Para determinados propósitos es útil incorporar el tratamiento lógico del conocimiento que se asume por parte de algunos grupos de agentes. Se usan entonces nuevos operadores, el de conocimiento del grupo E, en realidad, la abreviatura de que cada agente del grupo conoce la fórmula de que se trate, el de conocimiento común C, iteración *ad infinitum* del precedente, y el de conocimiento distribuido D, tal que si, por ejemplo, un agente a conoce $\varphi \rightarrow \psi$ y otro b conoce φ, se puede decir que el conjunto {a, b} conoce ψ. Omitimos los detalles de la semántica de estos operadores en aras de la brevedad.

Se han presentado diversos sistemas de creencias, aunque la idea subyacente es que una proposición creída puede ser verdadera, aunque la proposición como tal sea falsa. Es razonable la afirmación "Juan cree que es martes, aunque de hecho es miércoles", mientras que, tal como hemos usado "conocimiento" más arriba, sería absurdo afirmar "Juan sabe que es martes, aunque de hecho es miércoles" (van Ditmarsch, van der Hoek y Kooi 2008, 38). Para representar las habilidades de los agentes respecto de creencias —aquí haremos uso de ello para el abordaje lógico de la abducción—, se amplía el lenguaje ya conocido con un nuevo operador modal, B, con lo que $B_a\varphi$ representa que "el agente a cree que φ". Los marcos de Kripke para la interpretación del operador contendrán relaciones $\mathfrak{R}_{Ka}$, $\mathfrak{R}_{Ba} \subseteq W \times W$, para el conjunto de mundos W, de manera que las $\mathfrak{R}_{Ka}$ son las relaciones de accesibilidad respecto del conocimiento, que serán de equivalencia (reflexivas, simétricas y transitivas), mientras que las $\mathfrak{R}_{Ba}$ son relaciones de accesibilidad respecto de la creencia —en ambos casos, para el agente a—. Por otra parte, para cada agente a, $\mathfrak{R}_{Ba} \subset \mathfrak{R}_{Ka}$. Aunque se considere que lo conocido por un agente es creído por él, la afirmación recíproca sería falsa.

El sistema básico KD45 consta de los axiomas antes indicados, aunque cambiando el operador de conocimiento por el de creencia, además de sustituir el axioma 3 por el conocido como axioma D. Es decir, los axiomas son:

1. Todas las instancias de las tautologías proposicionales,

2. $K_a (\varphi \rightarrow \psi) \rightarrow (K_a\varphi \rightarrow K_a\psi)$ Axioma K,
3. $\neg B_a\bot$ Axioma D —consistencia de creencias—,
4. $B_a\varphi \rightarrow B_aB_a\varphi$ introspección positiva,
5. $\neg B_a\varphi \rightarrow B_a\neg B_a\varphi$ introspección negativa.

Aunque se mantienen las reglas de *modus ponens* y necesitación, esta última es relativa al operador de creencia, de φ se infiere $B_a\varphi$.

De cara a las aplicaciones epistemológicas —más abajo añadimos algunas observaciones sobre este punto—, mediante LED podemos abordar los problemas de los que ha ocupado el modelo AGM de revisión de creencias (elaborado en Alchourrón, Gärdenfors y Makinson 1985; véase también Schurz 2011), que contempla tres operaciones epistémicas, a saber, *expansión*, *contracción* y *revisión*. Se parte de un conjunto $\mathcal{K}$ de fórmulas del correspondiente lenguaje, una base de conocimiento, el cual es cerrado bajo consecuencia lógica (en sentido clásico), es decir, $\mathcal{K} = Cn(\mathcal{K})$. En esta breve presentación seguimos fundamentalmente el resumen expuesto en van Ditmarsch y Nepomuceno (2019). La expansión de $\mathcal{K}$ con una fórmula φ, en símbolos $\mathcal{K} + \varphi$ es el conjunto más pequeño de fórmulas que verifica:

1. $\mathcal{K} + \varphi$ es una base de conocimiento (cerrado bajo consecuencia).
2. $\varphi \in \mathcal{K} + \varphi$.
3. $\mathcal{K} \subseteq \mathcal{K} + \varphi$.
4. Si $\varphi \in \mathcal{K}$, entonces $\mathcal{K} + \varphi = \mathcal{K}$.
5. Si $\mathcal{K}' \subseteq \mathcal{K}$, entonces $\mathcal{K}' + \varphi \subseteq \mathcal{K} + \varphi$.

Por otra parte, la contracción de $\mathcal{K}$ con la fórmula φ, en símbolos $\mathcal{K} - \varphi$, da lugar a un conjunto de fórmulas que ha de verificar los siguientes postulados —solamente prestamos atención a los que han sido ampliamente aceptados, pues algunos de los iniciales opuestos en el modelo AGM han sido muy discutidos—,

1. $\mathcal{K} - \varphi$ es una base de conocimiento.
2. $\mathcal{K} - \varphi \subseteq \mathcal{K}$.
3. Si $\varphi \notin \mathcal{K}$ (o bien $\mathcal{K} \nvdash \varphi$), entonces $\mathcal{K} - \varphi = \mathcal{K}$.
4. Si $\varphi \in \mathcal{K}$, entonces $\mathcal{K} \subseteq (\mathcal{K} - \varphi) + \varphi$.
5. Si $\vDash \varphi \leftrightarrow \psi$, entonces $\mathcal{K} - \varphi = \mathcal{K} - \psi$.

La revisión de una base de conocimiento $\mathcal{K}$ mediante la fórmula φ, en símbolos $\mathcal{K} * \varphi$ se define mediante la identificación de Levi, como una contracción seguida de una expansión; en concreto,

$$\mathcal{K} * \varphi = (\mathcal{K} - \neg\varphi) + \varphi.$$

4. Sobre lógica abductiva

Se han presentado diversas formas de tratamiento lógico de la abducción, aunque se pueden clasificar en dos grupos: el clásico, denominado modelo AKM —de Aliseda, Kakas-Kowalski y Meheus-Maganani— y el conocido esquema GW —Gabbay y Woods—. En Aliseda (2006) y en Soler (2012) se define la abducción como se concibe en el modelo AKM, se trata de un proceso que produce explicaciones abductivas específicas con cierta estructura inferencial. Un problema abductivo, a su vez "desencadenante" de la solución abductiva, se puede representar como la terna $(\Theta, \varphi, \vdash_X)$, donde Θ representa una teoría base o base de conocimiento, un conjunto de fórmulas de un lenguaje dado, φ representa el fenómeno a explicar, un hecho "sorprendente" (como originalmente lo denominó Peirce) en el sentido de que no queda explicado (inferencialmente) por la teoría, mientras que $\vdash_X$ se refiere a la lógica subyacente de la práctica científica de que se trate, es decir, el sistema lógico que rige la comunidad científica en cuyo seno se realiza la práctica en cuestión (Hernández y Nepomuceno 2011). Una vez que se obtiene una solución ψ, de la teoría y la solución se puede inferir —según la lógica subyacente— φ. Es decir, se verificará

$$\text{si } \Theta, \varphi \Rightarrow_{abd} \psi, \text{ entonces } \Theta, \psi \vdash_X \varphi,$$

donde $\Rightarrow_{abd}$ expresa la inferencia abductiva propiamente dicha. Ahora bien, para que la abducción sea explicativa se debe cumplir que $\Theta \nvdash_X \varphi$ y que $\psi \nvdash_X \varphi$, de esta manera se evitan las explicaciones triviales. Una propuesta de tratamiento de la abducción en este marco mediante GTS lo hallamos en Hernández y Nepomuceno (2011).

En el planteamiento del esquema GW —sucintamente presentado en Barés-Gómez y Fontaine (2017)—, la abducción es una forma de inferencia basada en agentes y orientada hacia ciertos objetivos y no tiene por qué concluir en una proposición o en nuevo conocimiento, incluso podría no ser explicativa. Se alcanza una hipótesis que se toma como base para nuevas acciones, a pesar de que pueda persistir la ignorancia oficial; es una hipótesis que se puede utilizar y divulgar en otros razonamientos y acciones. El agente adopta la hipótesis obtenida mediante un razonamiento abductivo y queda momentáneamente como base para nuevas actuaciones cognitivas. Precisamente en Barés Gómez y Fontaine (2017) hallamos una aplicación de lógica dialógica para dar cuenta de desarrollos de la abducción tipo GW.

Se pueden considerar como complementarios ambos planteamientos. En cualquier caso, una certera caracterización de la abducción viene dada por las tesis de Kapitan-Hintikka (Hintikka 1998), que indicamos a continuación:

- Tesis inferencial. La abducción es, o incluye, un proceso inferencial, o varios procesos inferenciales.
- Tesis de propósito. En la investigación científica, el propósito de la abducción es doble: (i) generar nuevas hipótesis, y (ii) seleccionar hipótesis para su posterior verificación.
- Tesis de comprehensión. La abducción científica incluye todas las operaciones mediante las cuales se generan las teorías.
- Tesis de autonomía. La abducción es, o incorpora, un tipo de razonamiento que es distinto, e irreductible, a la inducción y a la deducción.

El fenómeno a explicar (el hecho sorprendente), desencadenante del proceso abductivo, si se presenta como una proposición (una fórmula del lenguaje de que se trate) y la hipótesis o conjetura alcanzada mediante abducción constituye el principal elemento explicativo, puede ser de dos clases, a saber, novedad o anomalía (Aliseda 2006). Simbólicamente, para el problema abductivo $(\Theta, \varphi, \vdash_X)$, tendremos:

1. Una novedad, si $\Theta \nvdash_X \varphi$ y $\Theta \nvdash_X \neg\varphi$.
2. Una anomalía, si $\Theta \nvdash_X \varphi$ pero $\Theta \vdash_X \neg\varphi$.

A partir de esta distinción es fácilmente equiparable la tarea de la búsqueda de soluciones a problemas abductivos, en el marco del modelo AKM de abducción, con un proceso de revisión de creencias, si bien considerando algunas restricciones para dar lugar a las nociones de AGM relativas a la abducción (Aliseda 2006). A este respecto, dado un problema abductivo $(\Theta, \varphi, \vdash_X)$, una solución abductiva ψ permite la explicación de φ, en el caso de ser una novedad, mediante una expansión abductiva, para abreviar, ExpAb $(\Theta, \varphi, \vdash_X)$, como

$$\text{ExpAb}\ (\Theta, \varphi, \vdash_X) = \{\psi \in L\ (\Theta) : \Theta, \psi \vdash_X \varphi\},$$

donde L (Θ) representa el lenguaje de la teoría base Θ y se cumplen las restricciones de la abducción explicativa. En el caso de la contracción, cuando φ representa una anomalía —$\Theta \vdash_X \neg\varphi$—, sea el conjunto $\{\beta_1,\ldots,\beta_n\} \subset \Theta$ tal que $\Theta\backslash\{\beta_1,\ldots,\beta_n\} \nvdash_X \neg\varphi$. De este modo,

$$\text{ContrAb}_{\neg\varphi} = \Theta \setminus \{\beta_1,\ldots,\beta_n\}.$$

Puesto que ante una anomalía se necesita una revisión, adaptamos la identidad de Levi, de manera que

$$\text{RevAb}\ (\Theta, \varphi, \vdash_X) = \text{ExpAb}\ ((\text{ContrAb}_{\neg\varphi}), \varphi, \vdash_X).$$

Desde el punto de vista de la LED es factible redefinir el lenguaje con las operaciones epistémicas de expansión, contracción y revisión, las cuales se representan para una fórmula φ, respectivamente, por +φ, -φ y ∗φ, por lo que las expresiones [+φ]ψ, [-φ]ψ y [∗φ] ψ deben leerse como "tras cada expansión con φ, ψ es el caso", "tras cada contracción con φ, ψ es el caso" y "tras cada revisión con φ, ψ es el caso", respectivamente. Teniendo en cuenta una vez más la identidad de Levi, cabe entender $[*\varphi]\psi = [-\neg\varphi]\,[+\varphi]\psi$. Los correspondientes operadores duales, ⟨+φ⟩, ⟨-φ⟩ y ⟨∗φ⟩, se definen como es habitual.

En cuanto a la semántica de los nuevos operadores, se han de tener en cuenta las acciones mismas. Para un modelo $M = (W, \Re_a, v)$ y estado s,

$$M, s \vDash [+\varphi]\psi \text{ syss si } M, s \vDash \varphi, \text{ entonces } M|_{+\varphi}, s \vDash \psi,$$

para $M|_{+\varphi} = (W^{+\varphi}, \Re^{+\varphi}_a, v^{+\varphi})$, $W^{+\varphi} = W \cup W'$, tal que para todo $s' \in W'$, $s' \in v^{+\varphi}(\varphi)$ y, para toda variable proposicional p, $v^{+\varphi}(p) = v(p)$ si p no ocurre en φ (en otro caso $v^{+\varphi}(p) = v(p) \cup W'$). Es decir, se trata de acciones epistémicas que permiten una actualización de la información en las que determinadas fórmulas valen. De manera análoga se ha de considerar la contracción:

$$M, s \vDash [-\varphi]\psi \text{ syss } M, s \vDash \varphi, \text{ entonces } M|_{-\varphi}, s \vDash \psi.$$

Ahora $W^{-\varphi} = W\backslash\{\beta_1,\ldots,\beta_n\}$, siempre que $W\backslash\{\beta_1,\ldots,\beta_n\}$ sea el mínimo conjunto tal que $W\backslash\{\beta 1,\ldots,\beta n\} \nvDash \varphi$. Por lo que respecta a la revisión, tendremos

$$M, s \vDash [*\varphi]\psi \text{ syss } M, s \vDash [-\neg\varphi][+\varphi]\psi.$$

Con estos nuevos operadores las nociones de expansión, contracción y revisión abductivas se pueden presentar en el siguiente formato. Sea M_Θ un modelo de la teoría Θ. Entonces, para el problema abductivo correspondiente,

- $\text{ExpAb}(\Theta, \varphi, \vdash_X) = \{\psi \in L(\Theta) : M_\Theta, s \vDash [+\psi]\varphi\}$, si φ es una novedad
- $\text{RevAb}(\Theta, \varphi, \vdash_X) = \{\psi \in L(\Theta) : M_\Theta, s \vDash [-\neg\varphi][+\psi]\varphi\}$, en el caso de que φ sea una anomalía.

5. Una anomalía: la lengua pirahã

En este apartado vamos a aplicar el aparato teórico expuesto anteriormente a un debate abierto en el ámbito de la lingüística, con motivo de presentar un proceso de razonamiento abductivo en una investigación científica concreta. Antes de comenzar, vamos a asumir, como presupuesto básico, que el paradigma actual de la gramática generativa es la teoría de Chomsky T(Ch).

La teoría de la gramática generativa defiende que existe una estructura del lenguaje universal, que nos permite, desde que somos pequeños, adquirir el lenguaje. En este senti-

do, la teoría defiende que todos los seres humanos tenemos la facultad innata del lenguaje, con lo que el lenguaje no sería una construcción humana desde cero, un artefacto humano, sino el desarrollo de un hipotético "órgano del lenguaje", que estaría ubicado en alguna parte del cerebro. Para reforzar nuestra propuesta de estudiar un problema lingüístico como un problema abductivo, vamos a señalar lo siguiente: el propio Chomsky, en su defensa de un lenguaje universal innato, aludió a Peirce a la hora de defender que aprendemos el lenguaje gracias a la capacidad humana de elegir las mejores hipótesis, de entre las infinitas hipótesis posibles, acerca de un tema: "Abduction also finds a place in theories of language acquisition. Most prominently, Chomsky proposed that learning a language is a process of theory construction. A child 'abduces' the rules of grammar guided by her innate knowledge of language universals" (Aliseda 2006, 44).

De todo ello, por tanto, se deriva según Chomsky la existencia de una estructura gramática común a todas las lenguas, por lo que se podrían definir leyes fundamentales de la gramática que estuvieran presentes en cualquiera. Una de las leyes o principios que Chomsky estableció como universal a toda lengua fue el principio de la *recursividad*, que consiste en la capacidad de poner una frase dentro de otra en una cadena infinita, a partir de las herramientas finitas del lenguaje.

No obstante, el lingüista y antropólogo, Daniel L. Everett, presentó, como crítica a la teoría chomskiana, el caso de la lengua de una tribu amazónica: la lengua pirahã. El estudio de Everett de la lengua pirahã establece que la gramática de esta se rige por un principio propio de la cultura pirahã, a saber, el principio de *inmediatez de la experiencia*. Según este principio un miembro de la tribu solo habla acerca de aquello de lo que tiene una experiencia directa o de testimonios comunicados por alguien que puede corroborar lo que el otro dice: "Los enunciados pirahã contienen únicamente afirmaciones directamente relacionadas con el momento en que se habla, tanto si se trata de una experiencia personal del hablante como de un hecho presenciado por un contemporáneo del hablante" (Everett 2014, 165). Este principio rector de las comunicaciones se ve reflejado en la gramática, según Everett, mediante la ausencia de cláusulas recursivas y oraciones subordinadas. Además, los conceptos de los pirahã se relacionan con la experiencia inmediata, con lo que no aceptan ningún nivel de abstracción más allá del que supone la creación misma de los conceptos que poseen.

En último término, se podría decir que el lenguaje pirahã respondería perfectamente a la proposición de Wittgenstein, según la cual los límites de nuestro mundo son los límites de nuestro lenguaje. El lenguaje pirahã, como afirma Everett, no representa nada que no sea accesible a la experiencia y, con ello, tampoco representa nada que no sea accesible a la cultura pirahã. En definitiva, la cultura pirahã, en concreto el principio de inmediatez de la experiencia, determinaría el lenguaje pirahã, de tal forma que, según Everett, la teoría chomskiana de la gramática universal, sustentada en el principio de la recursividad, quedaría en entredicho. En palabras de Everett:

> La lengua es el resultado de las sinergias entre los valores de una sociedad, la teoría de la comunicación, la biología, la fisiología, la física (las limitaciones propias de nuestro cerebro y nuestra capacidad fonética) y el pensamiento humano. Y creo que esto también es válido para el motor del lenguaje: la gramática. (Everett 2014, 253)

Aplicando los instrumentos teóricos que hemos expuesto, nos encontramos ante un problema abductivo (T(Ch), φ_p, $\vdash_{TL}$), en el que, dada la teoría de Chomsky, T(Ch), la lógica subyacente de la teoría lingüística, $\vdash_{TL}$, el caso de la lengua pirahã constituye un fenómeno discrepante, φ_p, con respecto a la teoría, debido a la ausencia de cláusulas recursivas y al condicionamiento fuerte de la gramática pirahã por la cultura. En definitiva, tenemos que T(Ch) $\nvdash_{TL} \varphi_p$ y nuestro propósito a este respecto es encontrar una solución abductiva, ψ, al problema.

Después de la aparición del trabajo de Everett, se produjo dentro de la esfera de Chomsky un proceso que podríamos denominar de revisión parcial de creencias, ya que se produjo una adaptación o concreción de la teoría chomskiana para dar cuenta del fenómeno de la gramática pirahã. Dentro de este proceso, el artículo más destacado fue el de los lingüistas Andrew Nevins, David Pesetsky y Cilene Rodrigues (NPR), los cuales cuestionaron desde el primer momento la identificación del lenguaje pirahã como un fenómeno sorprendente o anómalo. El grupo NPR comienza su respuesta a Everett afirmando que dentro de la teoría de la gramática universal se asume el hecho de que la gramática se forma o moldea en parte por la experiencia (véase Nevins, Pesetsky y Rodrigues 2009, 359).

En primer lugar, el grupo NPR afirma que las características del pirahã, que para Everett eran excepcionales, son fenómenos ya identificados en otras lenguas, con lo que serían problemas ya resueltos por la teoría T(Ch) o en proyecto de ser resueltos. Una de las cuestiones que el grupo NPR considera ya examinada en otras lenguas, y estudiada dentro de la teoría, es la cuestión referida a la conexión entre cultura y gramática, que para Everett es central en su teoría. La postura que el grupo NPR defiende es que la teoría de Chomsky, T(Ch), es una teoría más abierta y amplia de lo que Everett considera. En relación con lo que acabamos de exponer, el grupo NPR alude a la regla estructural en la construcción de frases, introducida por Chomsky en 2002, conocida como la *regla Merge*. De este modo, la recursividad queda asociada a una regla estructural, cuya aplicación no tiene por qué ser necesaria, en vez de ser definida como un principio fundamental de la teoría de la gramática generativa:

> Merge takes two linguistic units as input and combines them to form a set (a phrase), in which one element is designated as the phrase's head. Two kinds of linguistic units may serve as input to Merge: (i) lexical items, and (ii) phrases formed by previous applications of Merge. Since Merge may take previous applications of Merge as input, the rule is recursive. Iterated Merge yields the full variety of phrase

> structures studied in syntactic research—structures composed of lexical items and phrases that were themselves produced by Merge. (Nevins, Pesetsky y Rodrigues 2009, 365)

Con todo ello el grupo NPR quiere dejar claro que la naturaleza de la gramática universal está sujeta a nuevas interpretaciones, y que hay debates abiertos acerca de su capacidad de explicar cada uno de los fenómenos lingüísticos. Como señala el grupo NPR, no hay un modelo general de la gramática universal (véase Nevins, Pesetsky y Rodrigues 2009, 357), por lo que no estamos ante un sistema de conocimientos organizado deductivamente, al modo de la geometría de Euclides, por ejemplo:

> As a logical matter, of course, it is possible that beliefs considered nonnegotiable will turn out to be false, and it is never good to be so rigid about one's expectations that it becomes impossible for a new discovery to offer the element of surprise. One might disagree with Everett's claims about universal grammar, reject the IEP [immediacy of experience principle] and the claimed link between culture and grammar, and still agree that Pirahã grammar presents us with just such a surprise (in which case, one might wish to rethink aspects of syntactic theory in light of the new results). (Nevins, Pesetsky y Rodrigues 2009, 359)

En este sentido, el grupo NPR utiliza una lógica subyacente, $\vdash_{TL}$, que responde a los criterios de una lógica abductiva. Además, el grupo NPR no niega la posibilidad de añadir nuevas hipótesis explicativas a la teoría de la gramática universal para explicar nuevos fenómenos lingüísticos. Sin embargo, lo que niega el grupo de lingüistas es que el lenguaje pirahã constituya un fenómeno sorprendente, φ_p, que requiera una explicación particular; más bien debería hablarse de la necesidad de aplicar de una manera diferente la regla Merge. En este sentido, dada la teoría de Chomsky de la gramática generativa, con la regla estructural Merge, $T(Ch)_M$, y la hipótesis asumida dentro de la teoría, según la cual la experiencia, así como la cultura, interviene en la configuración de la gramática, ψ, tenemos que la existencia del lenguaje pirahã, φ_p, es un fenómeno que, según NPR, entronca con lo anterior

$$T(Ch)_M, \psi \vdash_{TL} \varphi_p.$$

Por tanto, el fenómeno pirahã constituiría una novedad asumida por la teoría, teniendo en cuenta la hipótesis ψ que Everett no consideraba, es decir, $T(Ch)_M \nvdash_{TL} \varphi_p$ y $T(Ch)_M \nvdash_{TL} \neg\varphi_p$, que solo requeriría una nueva configuración de las reglas estructurales de la gramática universal.

No obstante, el hecho de que la aparición de los estudios de Everett acerca del pirahã, siga siendo fuente de debates y discusiones dentro de la lingüística nos lleva a la siguiente pregunta: ¿es la lengua pirahã efectivamente una novedad o es más bien una anomalía? En este trabajo defendemos que el caso de la lengua pirahã es una anomalía vista o estudiada desde la propia teoría paradigmática, hablando en términos kuhnianos, como un reto o

rompecabezas a partir del cual se pueden seguir dos caminos: llevar a cabo un proceso de expansión abductiva de la teoría —como el grupo NPR— o llevar a cabo un proceso de revisión abductiva de la teoría chomskiana, que sería un camino más afín al proyecto de Everett. Por tanto, podríamos concluir que, dado que existe todavía un debate abierto, generado por el estudio de Everett, acerca de la capacidad de la teoría chomskiana de explicar fenómenos lingüísticos que no se infieren directamente de la teoría, como es el caso del pirahã, no podemos más que seguir considerando el pirahã como una anomalía, que no invalida la teoría chomskiana, pero que tampoco consigue hacer que lo anómalo se vuelva esperado:

$$T(Ch)_M \nvdash_{TL} \varphi_p \text{ pero } T(Ch)_M \vdash_{TL} \neg\varphi_p.$$

6. Conclusiones y trabajo futuro

Tras el giro dinámico en la lógica contemporánea vemos que en la investigación científica no podemos considerar como relevante la lógica clásica en exclusiva. La riqueza conceptual que se alcanza con el desarrollo de las lógicas no clásicas constituye un acicate para nuevas perspectivas en estudios epistemológicos. Como columnas en las que se sustenta plenamente la consumación del giro dinámico, la IF-logic (y semántica GTS), la lógica dialógica y el programa de dinámica lógica de la información, con los desarrollos de la LED como su expresión fundamental, ya han sido usadas en el estudio de la abducción, respectivamente en Barés Gómez y Fontaine (2017), Hernández y Nepomuceno (2011), y en Nepomuceno, Soler y Velázquez (2017) —en este mismo trabajo hemos explorado más de una línea de abordaje de la abducción con herramientas de LED—.

El sistema básico de LED que hemos presentado nos puede servir como una herramienta inicial para la consideración de una lógica subyacente en ciertas prácticas científicas. Por práctica científica hemos de entender la ejecución de un programa de investigación que realiza una comunidad científica, la cual se puede representar como un conjunto de agentes, en terminología de los sistemas multiagente. Por lógica subyacente debemos entender el conjunto de normas que rigen los modos de inferir de la comunidad en la práctica de que se trate. De esta manera, en cierto sentido, el sistema inferencial de la comunidad científica que lleva a cabo la práctica en cuestión es representable como una lógica. La continuidad de estas aproximaciones y aplicaciones deberá dirigir nuevas exploraciones, por ejemplo, con anuncios, y otras acciones epistémicas, realizadas por agentes concretos para grupos (no necesariamente el anuncio público como tal, sino dirigido a una clase de agentes). Asimismo, la consideración del tiempo, es decir, el uso de sistemas de lógica epistémica temporal.

Tanto los sistemas multiagente como, por ejemplo, el modelo AGM de revisión de creencias o sus variantes (véase, por ejemplo, Luna 2001), para estudios epistemológicos, se toman en préstamo de temáticas que fueron en su momento estudiadas en el ámbito de la Inteligencia Artificial (IA). Ello se justica por la representatividad que cabe atribuir a la

IA respecto de la historia de la ciencia en el pensamiento occidental. Si, como creemos, en la historia de la IA se condensan las principales problemáticas surgidas en la historia de la ciencia occidental, entonces cabe considerar la primera como un marco de referencia de estudios epistemológicos. Se trata de una cuestión de economía y eficiencia, ya que la IA tiene una fecha de nacimiento, 1956, con motivo de la famosa Conferencia de la Universidad de Dartmouth, y sesenta y tres años son más fáciles de abarcar que más de veinte siglos. En trabajos futuros se pueden explorar casos concretos de la historia de la IA calificables como de revisión de creencias. A este respecto, los sistemas de lógica defectuosa y su uso en el estudio de la abducción constituyen una línea abierta de indudable interés epistemológico.

En algunos de los usos de los sistemas estudiados se nos plantea un importante problema, el de la omnisciencia lógica. En el modelo AGM el agente haría las operaciones epistémicas a partir de bases de conocimiento cerradas bajo consecuencia, es decir, se trataría de un agente que es lógicamente omnisciente. Pero los agentes reales, los miembros de la comunidad científica, no lo son. En este caso, la propia LED nos permite extraer más herramientas conceptuales. Si un agente conoce una regla de inferencia R que, en general, podemos enunciar como "de *precondición*-R se infiere *postcondición*-R", entonces cabe pensar que el agente (no omnisciente) alcanza la *postcondición*-R siempre que, a partir de *precondición*-R, realice determinada acción. Sea ?R la acción de comprobar y asumir la regla de inferencia indicada, es decir, aprenderla; entonces se añadirán expresiones para acciones α —la más básica, aprender una regla, concatenación de acciones, etc.—, de manera que $[\alpha]\varphi$ expresará que "tras la acción α, φ es el caso" (una presentación concreta de un sistema de lógica de acciones aparece en van Ditmarsch, van der Hoek y Kooi 2008, 190 y ss.). A partir de aquí otra línea de trabajo debería conectar toda esta temática de los modelos de acción con el tratamiento lógico de la abducción. En conexión con ello, investigar las posibilidades de desarrollo de diálogos lógicos de carácter colaborativo y revisable (con posibles aplicaciones en tecnologías del lenguaje).

El sistema con operadores de expansión, contracción y revisión no calca plenamente, por así decir, las características de las correspondientes operaciones del modelo AGM. No obstante, en este contexto, precisamente el problema de la omnisciencia lógica recomienda un tratamiento diferente del original. Lo que nos interesan como punto de partida son las teorías base equiparables con teorías científicas, no ya cerradas bajo consecuencia, sino simplemente consistentes (lo que no excluye que en algunos contextos interesen sistemas paraconsistentes, lo que hemos excluido en este trabajo). Si bien las operaciones de expansión, contracción y revisión en AGM se rigen por unos postulados específicos sobre conjuntos de fórmulas, los operadores $+\varphi$ y $-\varphi$ y, por ende, $*\varphi$, que hemos incorporado en LED actúan sobre estructuras kripkeanas. Como quiera que no parece posible hallar una solución inequívoca al problema de la omnisciencia lógica, tal vez lo más razonable sea abordar las situaciones prácticas mediante la elección de los marcos y modelos disponibles que más o menos capten el grado de no-omnisciencia lógica aceptable en esa situación particular (Meyer, van der Hoek 2004, 89). Así, en la búsqueda de soluciones a

problemas abductivos (en el marco AKM) contamos con los sistemas comentados (LED con los diez esquemas axiomáticos, KD45, ampliación con acciones epistémicas, etc.). A este respecto, el sistema KL (Kraus y Lehmann, estudiado en Meyer, van der Hoek 2004) ofrece un intento de combinar conocimiento y creencia de una manera lo más realista posible y señala un nuevo sendero de trabajos futuros.

Hemos presentado de manera resumida el caso de la lengua pirahã, en el que se suscitan ciertos problemas epistemológicos de los que nos ocupamos en el trabajo. Su aparición como hecho sorprendente nos lleva a preguntarnos si hemos de considerarlo como una novedad y, en consecuencia, debemos buscar la solución como una expansión abductiva, o si se trata más bien de una anomalía, en cuyo caso se ha de proceder a una revisión de creencias. Tanto si se aborda el problema desde una perspectiva chomskiana o su opuesta, más afín a las conclusiones de Everett, entendemos que el enfoque correcto es el del tratamiento del hecho como una anomalía. A partir de aquí, cabe una más amplia y detallada explicación de este fenómeno en la que entrarán en juego buena parte de las expectativas antes referidas.

Referencias bibliográficas

Aho, T., Pietarinen, A. (eds.) (2006). *Truth and Games. Essays in Honor of Gabriel Sandu*. Acta Philosophica Fennica, Vol. 78. Helsinki: Societas Philosophica Fennica.

Alchourrón, C. E., Gärdenfors, P., Makinson, D. (1985). On the logic of theory change: partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50(2): 510-530. doi: 10.2307/2274239

Aliseda, A. (2006). *Abductive Reasoning. Logical Investigations into Discovery and Explanation*. Synthese Library, Vol. 330. Dordrecht: Springer.

Aliseda, A. (2014). *La Lógica como Herramienta de la Razón. Razonamiento Ampliativo en la Creatividad, la Cognición y la Inferencia*. Cuadernos de lógica, epistemología y lenguaje, Vol. 6. Londres: College Publications.

Baltag, A., Smets, S. (2012). The dynamic turn in quantum logic. *Synthese*, 186: 753-773. doi: 10.1007/s11229-011-9915-7

Barés Gómez, C., Fontaine, M. (2017). Argumentation and Abduction in Dialogical Logic. In L. Magnani, T. Bertolotti (eds.), *Springer Handbook of Model-Based Science*, 295-314. Dordrecht: Springer.

van Benthem, J. (2011). *Logical Dynamics of Information and Interaction*. Cambridge: Cambridge University Press.

van Ditmarsch, H., van der Hoek, W., Kooi, B. (2008). *Dynamic Epistemic Logic*. Synthese Library, Vol. 337. Dordrecht: Springer. doi: 10.1007/978-1-4020-5839-4

van Ditmarsch, H., Nepomuceno, A. (2019). Public announcements, belief expansion and abduction. En O. Pombo, A. Pato y J. Redmond (eds.), *Epistemologia, Lógica e Linguagem*, pp. 151-161. Lisboa: Centro de Filosofia das Ciências da Universidade de Lisboa.

Everett, D. L. (2014). *'No duermas, hay serpientes' Vida y lenguaje en la Amazonia*. Madrid: Turner Publicaciones S.L.

Fagin, R., Halpern, J. Y., Moses, Y., Vardi, M. Y. (1995). *Reasoning about Knowledge*. Cambridge, London: The M. I. T. Press.

Gochet, P. (2002). The dynamic turn in century logic. *Synthese*, 130: 175-184. doi: 10.1023/A:1014435213120

Hernández, I., Nepomuceno, A. (2011). Lógicas alternativas subyacentes a las prácticas científicas. En O. Pombo (ed.), *Lógica Universal e Unidade da Ciencia*, pp. 121-140. Lisboa: Centro de Filosofia das Ciências da Universidade de Lisboa.

Hintikka, J. (1973). *Logic, Language-Games and Information*. Oxford: Clarendon Press.

Hintikka, J. (1997). A revolution in the foundation of mathematics? *Synthese*, 111: 155-170. doi: 10.1023/A:1004970403579

Hintikka, J., Sandu, G. (1997). Game-Theoretical Semantics. In J. van Benthem and A. ter Meulen (eds.), *Handbook of Logic and Language*, pp. 361-410. Amsterdam: Elsevier.

Hintikka, J. (1998). What is abduction? The fundamental problem of contemporary epistemology. *Transactions of The Charles S. Peirce Society*, 34(3): 503-533. doi: 10.1007/978-94-015-9313-7_4

Luna, C. D. (2001). Una Generalización del Modelo AGM de Cambio de Creencias. *Inteligencia Artificial: Revista Iberoamericana de Inteligencia Artificial*, 5(13): 23-32.

Marion, M. (2006). Hintikka on Wittgenstein: From Language-Game to Game Semantics. *Acta Philosophica Fennica*, 78: 255-274.

Meyer, J. J. Ch., van der Hoek, W. (2004). *Epistemic Logic for AI and Computer Science*. Cambridge: Cambridge University Press.

Nepomuceno, A., Soler, F., Velázquez, F. R. (2017). Abductive Reasoning in Dynamic Epistemic Logic. In L. Magnani and T. Bertolotti (eds.), *Springer Handbook of Model-Based Science*, pp. 269-293. Dordrecht: Springer, doi: 10.1007/978-3-319-30526-4-13

Nevins, A., Pesetsky, D., Rodrigues, C. (2009). Pirahã exceptionality: A reassessment. *Language*, 2(85): 355-404.

Popper, K. R. (1980). *La lógica de la investigación científica*. Madrid: Tecnos.

Rahman, S., Clerbout, N., Keiff, L. (2009). Dialogues and Natural Deduction. In G. Primiero (ed.), *Acts of Knowledge, History, Philosophy, Logic*, pp. 301-336. London: College Publication.

Redmond, J. and M. Fontaine (2011). *How to Play Dialogues. An Introduction to Dialogical Logic*. Londres: College Publication.

Saarinen, E. (ed.) (1979). *Game-Theoretical Semantics. Essays on Semantics by Hintikka, Carlson, Peacocke, Rantala, and Saarinen*. Dordrecht: Springer.

Schurz, G. (2011). Abductive Belief Revision in Science. In E. J. Olsson and S. Enqvist (eds.), *Belief Revision meets Philosophy of Science*, pp. 77-104. Dordrecht: Springer. doi: 10.1007/978-90-481-9609-8-4

Soler, F. (2012). *Razonamiento abductivo en lógica clásica*. Londres: College Publications.

Revista de Humanidades de Valparaíso
Año 7, 2019, 1er semestre, No 13, págs. 90-101
eISSN 0719-4242

Humanities Journal of Valparaiso
No 13 (2019): 90-101

Paraconsistencia total

Total Paraconsistency

Bruno Da Ré

CONICET/Universidad de Buenos Aires, Argentina
bruno.horacio.da.re@gmail.com

DOI: https://doi.org/10.22370/rhv2019iss13pp90-101

Recibido: 27/06/2019. Aceptado: 16/08/2019

Resumen

Dentro del conjunto de las lógicas no clásicas, las lógicas paraconsistentes han suscitado de manera particular el interés de diversos filósofos. Además de las definiciones tradicionales, en los últimos años, se han propuesto nuevas maneras de caracterizar a la paraconsistencia. Lo que tienen en común todas estas definiciones es que alguna forma de la regla o de la metarregla de explosión debe ser rechazada. En este artículo, presentaré dichas definiciones y evaluaré el rol que juegan la negación y la transitividad en ellas. Finalmente, propondré una nueva caracterización de la paraconsistencia, la que llamaré *paraconsistencia total* y mostraré que la regla de monotonía juega un rol crucial en todas las presentaciones de dicho concepto.

Palabras clave: negación, transitividad, monotonía, lógicas paraconsistentes, explosión.

Abstract

In the context of non-classical logics, many philosophers have been particularly interested in the paraconsistent logics. In addition to traditional definitions, in recent years, new ways of characterizing the notion of paraconsistency have been proposed. In all of these definitions the rule or the meta-rule of explosion is abandoned. In this article, I present those definitions and evaluate the role that the negation and the transitivity play in each of them. Finally, I propose a new definition of paraconsistency which I call *total paraconsistency* and show that the rule of weakening plays a crucial role in all of the characterizations of such a concept.

Keywords: negation, transitivity, weakening, paraconsistent logics, explosion.

1. Introducción

La noción de paraconsistencia y especialmente sus motivaciones filosóficas han sido tematizadas y revisadas en numerosas ocasiones y con diferentes propósitos. Sin embargo, desde el punto de vista conceptual, aún se mantiene vigente la clasificación elaborada por Igor Urbas, en su célebre artículo *Paraconsistency* (Urbas 1990). Allí, el autor realiza una separación entre tres proyectos, aparentemente disjuntos, que han motivado el desarrollo de lógicas paraconsistentes.

En primer lugar, la *posición dialeteica* (o tradición australiana-americana) consiste en aquellos que creen o sostienen que hay contradicciones verdaderas. Por supuesto, esto no significa que toda contradicción sea verdadera (esto sería una posición trivialista). En esta línea de análisis, podemos encontrar a todos aquellos filósofos y lógicos que han intentado desarrollar teorías paraconsistentes que den cuenta de fenómenos paradójicos. Entre otros, podemos mencionar a Priest (1979; 2006), Routley y Meyer (1976), entre otros.

En segundo lugar, se encuentra la *posición pragmática*. Dentro de esta línea estarían quienes simplemente advierten que en algunos contextos hay proposiciones contradictorias, por ejemplo en códigos legales, en nuestras creencias, en teorías científicas, y no por eso de esa contradicción debería seguirse cualquier cosa. Así, la paraconsistencia serviría como base para expresar teorías inconsistentes. En esta línea, se encuentra la tradición brasileña y cabe mencionar a su fundador, Da Costa (1974) y en la actualidad a Carnielli y Coniglio (2016).

Por último, Urbas menciona las llamadas *posiciones independientes*. Entre estos se encuentran quienes no necesariamente se comprometen con la existencia de contradicciones verdaderas ni con la postulación de teorías inconsistentes, sino que restringen por otras razones la aplicación de la regla de explosión. Entre ellos se encuentra toda la tradición de lógicos relevantes quienes, a los fines de hacer válida una inferencia, prescriben la necesidad de algún tipo de conexión entre premisas y conclusión (Mares y Meyer 2001).

Sin embargo, ¿qué significa que una lógica sea paraconsistente? Dependiendo la posición que cada filósofo y lógico ha tomado, la paraconsistencia ha sido definida de distintas maneras[1]. Al menos hay dos definiciones disponibles de paraconsistencia en la literatura. Por un lado, tenemos la célebre definición de Da Costa (1974, 498), quien afirmó que toda lógica paraconsistente debe satisfacer los siguientes requisitos[2]:

1. El principio de no contradicción $\neg (A \wedge \neg A)$ no debe ser un teorema.

[1]No es la intención de este artículo evaluar la plausibilidad de la adopción de una lógica paraconsistente. Un análisis detallado respecto de la relación entre la noción técnica de paraconsistencia y sus implicaciones filosóficas puede encontrarse en Barrio y Da Re (2018) o Barrio (2018). En particular, no evaluaré los posibles problemas asociados a tales lógicas. Para ello, puede consultarse v.g. Bobenrieth (1998).

[2]Si bien aquí hemos citado la definición más célebre, otras definiciones similares, tal vez más detalladas, pueden encontrarse en da Costa y Lewin (1995).

2. De dos fórmulas contradictorias no debe seguirse una fórmula arbitraria.
3. La lógica debe tener una extensión simple de primer orden.
4. La lógica debe ser tan clásica como sea posible, satisfaciendo las condiciones 1-3.

Como puede verse esta serie de condiciones son condiciones necesarias para que una lógica sea paraconsistente.

Sin embargo, dado que la paraconsistencia es una propiedad técnica que se aplica o no a una lógica, es preciso presentar definiciones más específicas acerca de este concepto. En este orden de ideas, en la bibliografía usual, se pueden encontrar las siguientes[3]:

> Definición 1. Dada una lógica L, decimos que es paraconsistente si y sólo si existen dos fórmulas del lenguaje A, B, tal que $A \wedge \neg A \nvdash B$, esto es, la regla de explosión expresada de esta manera no es derivable en la lógica.
>
> Definición 2. Dada una lógica L, decimos que es paraconsistente si y sólo si existen dos fórmulas del lenguaje A, B, tal que $A, \neg A \nvdash B$, esto es, la regla de explosión expresada de esta manera no es derivable en la lógica.

La única diferencia entre ambas definiciones se encuentra en relación con la forma de combinar premisas, esto es mediante comas o mediante conjunciones. En este sentido, la primera de estas definiciones caracteriza lo que se conoce como *paraconsistencia conjuntiva*, mientras que la segunda lo que se conoce como *paraconsistencia colectiva*.

Finalmente, en un artículo reciente, Barrio, Pailos y Szmuc (2018) argumentan que la definición de paraconsistencia debe ser diferente.

> Definición 3. Dada una lógica L, decimos que es paraconsistente si y sólo si existen dos fórmulas del lenguaje A, B, tal que:

- La regla de explosión no es derivable, esto es $A, \neg A \nvdash B$ o bien
- La metarregla de explosión no es derivable[4], donde dicha metarregla consiste en:

$$\frac{\Rightarrow A \qquad A \Rightarrow}{\Rightarrow B}$$

[3]Por ejemplo, definiciones similares a estas pueden encontrarse en el reciente libro Carnielli y Coniglio (2016).

[4]Cabe destacar que en rigor los autores requieren que la metarregla de explosión sea localmente inválida. De todas formas, en este contexto tomaremos esta definición, a los fines de mantener la coherencia entre todas las definiciones. Además, formulan dicha metarregla utilizando una negación. Al definir de esta manera la metarregla de explosión, aquí estamos haciendo hincapié en la noción de transitividad, como veremos en la sección 2.

Esta última definición es un poco más abarcativa que las dos anteriores, dado que hace mención no sólo a la regla de explosión sino también a la metarregla de explosión. Esquemáticamente, una metarregla es una regla que establece relaciones metainferenciales, es decir regula las relaciones lógicas entre las inferencias. Así como las reglas establecen las transiciones inferenciales (lógicas) entre fórmulas, las metarreglas las establecen entre las inferencias mismas. Por esa razón, la lógica determinada por las reglas suele llamarse *lógica interna*, mientras que la lógica de las metarreglas es la *lógica externa* (Avron 1991)

En las siguientes secciones utilizaré tres lógicas: LP, ST y RND que pueden ser consideradas paraconsistentes en alguno o en todos los sentidos mencionados[5]. En la sección 2, analizaré las definiciones que en esta sección se han presentado en relación con las primeras dos lógicas (ST y LP). En la sección 3, me centraré en la lógica RND e introduciré una definición de "paraconsistencia total" que mostrará que la nota característica de la paraconsistencia se encuentra íntimamente relacionada con la monotonía. Finalmente, en la sección 4 recopilaré las conclusiones de este análisis.

2. Dos lógicas parcialmente paraconsistentes

En las siguientes dos subsecciones nos ocuparemos de dos lógicas que podríamos caracterizar como parcialmente paraconsistentes: LP y ST. Esta caracterización será explicada más adelante, pero baste mencionar que se debe a que tanto LP como ST son lógicas que aceptan o derivan alguna forma explosión. En el caso de LP, en tanto regla; en el caso de ST, en tanto metarregla.

2.1 LP: modificando la negación

Esta lógica desde un punto de vista formal fue propuesta originalmente por Asenjo (1966), si bien fue Priest (1979) quien la denominó Logic of Paradox (LP) y quien ha desarrollado una prolífica producción alrededor de sus aplicaciones. Dado que en este artículo estamos interesados en la paraconsistencia, presentaremos el fragmento conjunción y negación de la lógica de manera semántica[6]. Sea L un lenguaje proposicional, V={0,i,1} el conjunto de valores semánticos y D={1, i} el conjunto de valores designados. Las conectivas se comportan de acuerdo con los esquemas de Strong Kleene que a continuación se detallan:

[5]Si bien como hemos mencionado no nos enfocaremos en las interpretaciones filosóficas de la paraconsistencia, cabe mencionar que, utilizando la clasificación que hemos elaborado al comienzo de esta sección, las primeras dos lógicas podrían ser caracterizadas como perteneciendo a la tradición dialeteica (si bien habría que hacer alguna salvedad en el caso de ST), mientras que RND pertenece a la tradición relevante, y por ende lo que llamamos posiciones independientes.

[6]Un cálculo de secuentes correcto y completo se puede encontrar en Beall (2011).

$f\ \neg$	
1	0
i	i
0	1

f $\wedge$	1	i	0
1	1	i	0
i	i	i	0
0	0	0	0

Una inferencia $\Gamma \models A$ es válida en LP si y sólo si, para toda valuación v, si $v(\gamma)=1$ o i para cada premisa, entonces $v(A)= 1$ o i. En otras palabras, validez en LP es definida como preservación de valor designado.

Ahora bien, LP es una lógica paraconsistente, dado que la regla de explosión no es válida: $A, \neg A \not\models B$. La valuación *v* que es un contraejemplo para dicha inferencia es aquella tal que: $v(A)= v(\neg A) = i$ y $v(B)=0$. Lo mismo sucede si formulamos la regla de explosión de manera conjuntiva. Como puede notarse, la razón por la que LP invalida la regla de explosión se debe a que la negación es tal que es posible que una fórmula y su negación reciban valor designado en alguna valuación. En otras palabras, la presencia de una negación no clásica (también llamada *débil*, en contraposición con la negación *fuerte* o clásica) es lo que permite la falla de explosión. Por otro lado, dado que la noción de consecuencia lógica de LP es una noción tarskiana, esto es transitiva, monotónica y reflexiva, la metarregla de explosión resultará derivable en su cálculo de secuentes. En la próxima sección, veremos el caso inverso. Esto es, una lógica tal que la regla de explosión es inválida, pero la metarregla de explosión no es derivable.

2.2 ST: un enfoque no transitivo

En la literatura reciente sobre paradojas semánticas, Cobreros et al. (2013) y Ripley (2013) han propuesto el abandono de la noción tarskiana de consecuencia lógica y, en este sentido, han desarrollado una lógica no transitiva: ST. La idea básica de la propuesta se basa en la lectura de la noción de consecuencia lógica en términos de *posiciones*. Todo par $<\Gamma, \Delta>$ de conjuntos de fórmulas es una posición tal que las premisas (Γ) son aserciones y las conclusiones (Δ) son rechazos. Así, se define una posición como *fuera de los límites* si es incoherente aseverar todas las premisas mientras se deniegan todas las conclusiones. El caso de las oraciones paradójicas es el siguiente. La semántica va a ser caracterizable en términos de modelos trivaluados, mientras que la representación del valor intermedio va a plasmarse en el concepto de tolerancia. Hay dos formas de aserción: aserción estricta y aserción tolerante. Las oraciones paradójicas son asertadas tolerantemente, no estrictamente. Obviamente, lo mismo se aplica para la negación de dichas oraciones. Entonces, en términos estrictos, las oraciones paradójicas no son verdaderas ni son falsas. En términos tolerantes, son verdaderas y falsas. Por ende, el valor intermedio con el que se evalúa en todo modelo a las oraciones paradójicas es interpretado en este sentido. En términos

más generales, si una oración recibe el valor 1 en un modelo, diremos que la oración es tolerante y estrictamente asertable. Por otro lado, si la oración recibe 0 en un modelo, diremos que la oración no es asertable ni tolerante ni estrictamente. Por último, en el caso de que dicha oración reciba el valor intermedio, diremos que la oración es tolerantemente pero no estrictamente asertable. Así, se distingue el acto de habla de asertar estrictamente como un acto de habla más fuerte que asertar tolerantemente. Esta distinción postulada en el nivel de la pragmática les permite a los autores caracterizar el estatus de las oraciones paradójicas y otras oraciones problemáticas (por ejemplo, oraciones que involucran predicados vagos). En palabras de Ripley (2013, 153-154):

A tolerant assertion of something is in bounds exactly when a strict denial of that thing is out of bounds, and a tolerant denial of something is in bounds exactly when a strict assertion of that thing is out of bounds. (…) If there were no gap between strict assertibility and strict deniability, if cut held, then strict and tolerant would coincide. As this is not the case, they do not. Some things (the paradoxes) can be neither strictly asserted nor strictly denied, and so they can be both tolerantly asserted and tolerantly denied.

De esta manera, queda asegurada la falla de transitividad. Formalmente, desde el punto de vista semántico, tomemos las matrices de Strong Kleene ya presentadas. La lógica ST queda definida:

> Definición 4: $\Gamma \vDash \Delta$ si y sólo si para toda valuación *v*, si v(A) = 1 para todo A en Γ, entonces v(B)= 1 o i, para algún B en Δ.

La noción de consecuencia lógica es no transitiva dado que, asumiendo que hay una oración que recibe el valor intermedio en toda interpretación (una oración paradójica L) tenemos que la siguiente metainferencia es inválida:

$$\Gamma \vDash L \text{ y } L \vDash \Delta, \text{ implica } \Gamma \vDash \Delta, \text{ para cualquier } \Gamma, \Delta.$$

La razón de ello es que, dado que la lógica no es trivial, $\Gamma \nvDash \Delta$ para cualquier Γ, Δ, pero sí $\Gamma \vDash L$ y $L \vDash \Delta$, ya que v(L) = i, para toda valuación. Desde el punto de vista de teoría de la prueba, la regla de corte no es admisible en este sistema (Ripley 2013).

Cabe destacar que, si bien la lógica es no transitiva, desde el punto de vista inferencial, ST coincide extensionalmente con la lógica clásica. En otras palabras, toda inferencia clásicamente válida será válida en ST. En este sentido, un corolario directo de esta observación es que ST no es paraconsistente de acuerdo con las Definiciones 1 y 2 mencionadas en la introducción del presente artículo. Esto es, la regla de explosión es válida en ST[7].

[7]Para profundizar en la discusión acerca de la relación específica entre ST y la lógica clásica, puede consultarse Barrio, Pailos y Szmuc (2019).

Sin embargo, en este contexto comienza a jugar un rol la Definición 3. La metarregla de explosión no es derivable en el cálculo que usualmente se presenta para ST. Debido a que la lógica es no transitiva, de tener una prueba de $\vDash L$ y $L \vDash$ no se sigue que podamos derivar cualquier inferencia.

Entonces, recapitulando, así como en el caso de LP teníamos una lógica en la que fallaba la regla de explosión, pero no la metarregla de explosión, en el caso de ST nos encontramos con el caso inverso: la regla de explosión es derivable, pero no así la metarregla. Esta falla de explosión en un nivel y no en el otro motiva la siguiente pregunta: ¿qué es lo característico de una lógica paraconsistente? A esta pregunta, responderemos en la siguiente sección, cuando presentemos la lógica no monotónica RND.

3. La lógica relevante RND y la monotonía

En esta sección presentaré la lógica no monotónica que usaremos para problematizar las definiciones de paraconsistencia presentadas en las secciones anteriores. En primer lugar, nuestro objetivo es presentar una lógica no monotónica, esto es, una lógica en la que una regla estructural no es admisible. La lógica de la que aquí nos ocuparemos es la célebre lógica relevante RND[8].

En términos muy generales, las lógicas relevantes surgen como reacción a las llamadas paradojas de la implicación material. Por citar algunas, podemos mencionar las siguientes: $A \rightarrow (B \rightarrow A)$, $\neg A \rightarrow (A \rightarrow B)$, $(A \rightarrow B) \vee (B \rightarrow C)$ y sus contrapartes en términos de consecuencia lógica. Los teóricos relevantes han resaltado lo poco intuitivo de estas inferencias y otras vinculadas con el condicional y con la relación de consecuencia lógica, poniendo de manifiesto la ausencia aparente de relevancia para la conclusión de las premisas, o del consecuente del antecedente.

Así, los teóricos de la relevancia en sus trabajos, entre los cuales se destacan Anderson y Belnap (1975), Plumwood et al. (1982), Mares (2004) y Mares y Meyer (2001), sólo por nombrar algunos, desde un punto de vista informal suelen afirmar que la noción de consecuencia lógica está vinculada a la relación de compartir contenido o tópico. Pareciera que las falacias o paradojas del condicional o de la implicación están vinculadas con el hecho de que las premisas y la conclusión no comparten ningún tipo de contenido[9]. Una de las maneras que han encontrado dichos teóricos para capturar esta noción informal de relevancia es lo que se conoce como *principio de variable compartida*. Básicamente, la

[8]Esta lógica es simplemente el fragmento no distributivo de la lógica R.

[9]Sería completamente injusto con la literatura sobre lógicas relevantes si no mencionara que existen distinciones, técnicas y filosóficas, entre cada una de las propuestas defendidas por los autores, así como una amplia variedad de lógicas consideradas relevantes. Basten, sin embargo, estos breves comentarios a los fines únicamente de introducir y presentar las motivaciones más importantes que engloban a las diferentes posiciones.

idea es que toda lógica que pretenda ser relevante al menos debe satisfacer que la propiedad que establece que las premisas y la conclusión comparten una letra proposicional. Si bien esto es necesario, no es suficiente como para caracterizar una lógica como relevante.

La lógica RND es un fragmento de la lógica R, una de las lógicas relevantes más estudiadas en la literatura. En lugar de seguir sus presentaciones axiomáticas usuales, haré una presentación en términos de cálculos de secuentes para el lenguaje $\{\wedge, \neg, \rightarrow\}$. La lógica RND queda definida de la siguiente manera[10]:

Axiomas

$$A \Rightarrow A \text{ (Id)}$$

Reglas estructurales

$$\frac{\Gamma \Rightarrow A, A, \Delta}{\Gamma \Rightarrow A, \Delta}\text{(CR)} \qquad \frac{\Gamma, A, A \Rightarrow \Delta}{A, \Gamma \Rightarrow \Delta}\text{(CL)}$$

$$\frac{\Gamma \Rightarrow A, \Delta \qquad A, \Sigma \Rightarrow \Pi}{\Sigma, \Gamma \Rightarrow \Delta, \Pi}\text{ (Corte)}$$

Reglas operacionales

$$\frac{\Gamma \Rightarrow A, \Delta}{\neg A, \Gamma \Rightarrow \Delta}(\neg\text{-L}) \qquad \frac{A, \Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta, \neg A}(\neg\text{-R})$$

$$\frac{\Gamma, A \Rightarrow \Delta}{A \wedge B, \Gamma \Rightarrow \Delta}(\wedge\text{-L}) \qquad \frac{B, \Gamma \Rightarrow \Delta}{A \wedge B, \Gamma \Rightarrow \Delta}(\wedge\text{-L})$$

$$\frac{\Gamma \Rightarrow A, \Delta \qquad \Gamma \Rightarrow \Delta, B}{\Gamma \Rightarrow \Delta, A \wedge B}\text{ (}\wedge\text{-R)}$$

$$\frac{\Gamma \Rightarrow A, \Delta \qquad B, \Sigma \Rightarrow \Pi}{A \rightarrow B, \Gamma, \Sigma \Rightarrow \Delta, \Pi}\text{ (}\rightarrow\text{-L)}$$

[10]Tomamos este cálculo de Paoli (2002).

$$\frac{A, \Gamma \Rightarrow B, \Delta}{\Gamma \Rightarrow A \rightarrow B, \Delta} \ (\rightarrow\text{-R})$$

donde Γ, Δ, Σ y Π son multiconjuntos[11] de fórmulas, mientras que A y B son fórmulas del lenguaje. Las reglas (C-L) y (C-R) son las reglas de contracción. Por otro lado, por cada conectivo (*) hay una regla de introducción a la izquierda (*-L) y una regla de introducción a la derecha (*-R). La noción de consecuencia lógica que se desprende del cálculo anterior es la siguiente: $\Gamma \Rightarrow \Delta$ si y sólo si la inferencia $\Gamma \vdash \Delta$ es válida en RND.

En términos de secuentes, la lógica así definida es no monotónica, dado que en el cálculo no se encuentran explícitas (ni son admisibles) las siguientes metarreglas:

$$\frac{\Gamma \Rightarrow \Delta}{\Gamma \Rightarrow A, \Delta} \ (\text{WR}) \qquad \frac{\Gamma \Rightarrow \Delta}{A, \Gamma \Rightarrow \Delta} \ (\text{WL})$$

En otras palabras, del hecho de que un secuente $\Gamma \Rightarrow \Delta$ sea derivable (y por ende válido), no se sigue que el secuente que resulta de agregarle premisas (o conclusiones) a $\Gamma \Rightarrow \Delta$ sea derivable (ni sea válido).

Gracias a lo antedicho, del cálculo anterior se desprende que la regla de explosión no es derivable en esta lógica. En términos técnicos, comenzando con identidad y aplicando (¬-L), podemos derivar el secuente $A, \neg A \Rightarrow$. Sin embargo, la ausencia de monotonía evita que agreguemos una fórmula en la conclusión, y por ende el siguiente secuente no es derivable $A, \neg A \Rightarrow B$. Volviendo a la discusión filosófica sobre relevancia, explosión puede ser vista como una inferencia irrelevante, dado que permite que la conclusión y las premisas contradictorias no compartan ninguna variable proposicional (de hecho, todas las lógicas relevantes rechazan explosión).

Por otro lado, cabe mencionar que la relación de consecuencia lógica definida por la lógica RND no es tarskiana, en tanto no es monotónica[12]. En este sentido, cabe preguntarse qué sucede con la metarregla de explosión.

Así, mientras que las lógicas LP y ST podrían ser caracterizadas como parcialmente paraconsistentes, en tanto ambas aceptan una forma de explosión (la primera como metarregla, la segunda como regla), queda abierta la posibilidad de definir lo que llamaremos *paraconsistencia total:*

> Definición 5: Dada una lógica L, decimos que es *totalmente paraconsistente* si y sólo si existen dos fórmulas del lenguaje A, B, tal que:

[11]Un multiconjunto es una colección de objetos sensible a la repetición de elementos. Por ejemplo, mientras {φ, φ, ψ} y {φ, ψ} son el mismo conjunto, son diferentes multiconjuntos (el primero tiene tres elementos, mientras que el segundo sólo dos).

[12]Esta generalización de la noción de consecuencia lógica, junto con una sistematización de otras posibilidades puede hallarse en Avron (1991).

- La regla de explosión no es derivable, esto es A, ¬A ⊮ B, y
- La metarregla de explosión no es derivable, donde dicha metarregla consiste en:

$$\frac{\Rightarrow A \qquad A \Rightarrow}{\Rightarrow B}$$

Entonces, las lógicas que rechacen la regla de explosión como LP serán *paraconsistentes en el nivel inferencial.* Por otro lado, las que hagan lo propio con la metarregla de explosión como ST serán *paraconsistentes en el nivel metainferencial.* Como hemos mencionado, en ambos casos, consideraremos que dichas lógicas son parcialmente paraconsistentes.

Por otro lado, la lógica RND no sólo rechaza la regla de explosión en el nivel inferencial, sino que además la metarregla de explosión tampoco es derivable. En términos técnicos, de agregar al cálculo dos secuentes del siguiente tipo: ⇒ A y A ⇒, aplicando la regla Corte, podemos derivar el secuente vacío: ⇒. Sin embargo, de dicho secuente no es posible agregar ni premisas ni conclusiones (dado que las únicas metarreglas que lo permitirían son (W-L) y (W-R), que están ausentes). Por lo tanto, la metarregla de explosión no es derivable. Esto es, dicha lógica rechaza toda forma de explosión, por lo que, de acuerdo con la definición, será considerada como *totalmente paraconsistente*.

Así, más allá de lo definicional, en la caracterización de una lógica como totalmente paraconsistente se pone en primer plano la falla de monotonía, mientras que las caracterizaciones parciales están más vinculadas con la negación o la transitividad. En algún sentido, estas consideraciones pueden ser vistas como una extensión del artículo de Urbas (1990). Allí, el autor mostró que el abandono de monotonía permite el desarrollo de lógicas paraconsistentes respecto de cualquier conectiva lógica (es decir, permite la postulación de lógicas para las que toda instancia de explosión es inválida). Aquí, he mostrado que algo similar sucede en el nivel de las metainferencias: el abandono de monotonía permite que la metarregla de explosión no sea derivable.

4. Conclusiones

En este artículo he tematizado las nociones usuales de paraconsistencia en relación con las lógicas LP, ST y RND. A la luz de dichas conceptualizaciones, he mostrado que tanto LP como ST son lógicas que podríamos caracterizar como parcialmente paraconsistentes, dado que ambas aceptan alguna forma de explosión.

Mientras LP rechaza la regla de explosión, acepta su contraparte metainferencial, dado que su noción de consecuencia lógica es tarskiana. Por otro lado, ST coincide en el nivel inferencial con la lógica clásica y por ende resulta válida la regla de explosión. Sin embargo, dado que su noción de consecuencia lógica es no transitiva, esto permite eludir la derivablilidad de la metarregla de explosión.

Por último, he propuesto una nueva caracterización de la noción de paraconsistencia, la que he llamado *paraconsistencia total*. Simplemente, lo que se exige para que una lógica sea totalmente paraconsistente es que rechace explosión, tanto como regla como como metarregla. Así, he mostrado que la lógica RND es totalmente paraconsistente. Esto revela un aspecto más general: la falla de monotonía permite la falla de explosión en el nivel inferencial y en el nivel metainferencial. En este sentido, de todo lo expuesto se desprende que la monotonía está íntimamente relacionada con la paraconsistencia, quizás más de lo que se suele considerar. Cabe destacar para concluir que, obviamente, no es mi intención afirmar o sostener que de hecho LP o ST no son paraconsistentes. Por el contrario, creo que lo son, sólo que parcialmente: el rechazo de toda forma de explosión requiere el abandono de monotonía.

Referencias bibliográficas

Anderson, A., Belnap, N. (1975). *Entailment*, Vol. I. Princeton: University Press. doi: 10.2307/2272137

Avron, A. (1991). Simple consequence relations. *Information and Computation*, 92(1): 105-139. doi: 10.1016/0890-5401(91)90023-U

Asenjo, F. G. (1966). A calculus of antinomies. *Notre Dame Journal of Formal Logic*, *7*(1): 103-105. doi:10.1305/ndjfl/1093958482

Barrio, E. (2018). Models & Proofs: LFIs without a Canonical Interpretation. *Principia*, 22(1): 87-112 doi: 10.5007/1808-1711.2018v22n1p87

Barrio, E., Pailos, F., Szmuc, D. (2018). What is a paraconsistent logic? En W. Carnielli, J. Malinowski (eds.), *Contradictions*, *from Consistency to Inconsistency*, pp. 89-108. Cham: Springer. doi: 10.1007/978-3-319-98797-2_5

Barrio, E., Pailos, F., Szmuc, D. (2019). A Hierarchy of Classical and Paraconsistent Logics. *Journal of Philosophical Logic*. doi.org/10.1007/s10992-019-09513-z

Beall, J. (2011). Multiple-conclusion LP and default classicality. *The Review of Symbolic Logic*, 4(2): 326-336. doi:10.1017/S1755020311000074

Bobenrieth, A. (1998). Five philosophical problems related to paraconsistent logic. *Logique et Analyse*, 41(161-163): 21-30.

Carnielli, W., Coniglio, M. (2016). *Paraconsistent logic: Consistency, contradiction and negation*. Switzerland: Springer International Publishing. doi: 10.1007/978-3-319-33205-5

Cobreros, P., Égré, P., Ripley, D., Van Rooij, R. (2013). Reaching transparent truth. *Mind*, 122(488): 841-866. doi: 10.1093/mind/fzt110

da Costa, N. (1974). On the theory of inconsistent formal systems. *Notre Dame Journal of Formal Logic*, 15: 497-510. doi: 10.1305/ndjfl/1093891487

da Costa, N., Lewin, R. (1995). Lógica paraconsistente. En C. Alchourrón, J. Méndez, R. Orayen (eds.). *Lógica*. Enciclopedia IberoAmericana de Filosofía, Vol. 7, pp. 185-204. Madrid: Trotta.

Mares, E., Meyer, R. (2001). Relevant Logics. En Goble, Lou (ed.), *The Blackwell Guide to Philosophical Logic*. Oxford: Blackwell.

Mares, E. (2004). *Relevant Logic: a Philosophical Interpretation.* Cambridge: Cambridge University Press. doi:10.1017/CBO9780511520006

Paoli, F. (2002). *Substructural logics: a primer*. Dortrech: Springer Netherlands. doi: 10.1007/978-94-017-3179-9

Plumwood, V., Routley, R., Meyer, R., Brady, R. (1982). *Relevant Logics and its Rivals*, Volume I. Atascardero, CA: Ridgeview. doi: https://doi.org/10.2307/2275039

Priest, G. (1979). The logic of paradox. *Journal of Philosophical logic*, 8(1): 219–241. doi:10.1007/BF00258428

Priest, G. (2006). *In Contradiction: A Study of the Transconsistent.* Oxford: Oxford University Press. doi: 10.2307/2219835

Ripley, D. (2013). Revising up: Strengthening classical logic in the face of paradox. *Philosophers Imprint,* 13(5): 1-13. url: http://hdl.handle.net/2027/spo.3521354.0013.005

Routley, R., Meyer, R. (1976). Dialectical logic, classical logic, and the consistency of the world. S*tudies in East European Thought*, 16(1): 1-25. doi: 10.1007/BF00832085

Urbas, I. (1990). Paraconsistency. *Studies in Soviet Thought*, 39(3-4): 343-354. doi: 10.1007/BF00838045

Revista de Humanidades de Valparaíso
Año 7, 2019, 1er semestre, No 13, págs. 102-108
eISSN 0719-4242

Humanities Journal of Valparaiso
No 13 (2019): 102-108

– Reseña de libro –

La subversión de la democracia

Steven Levitsky y Daniel Ziblatt (2018). *Cómo mueren las democracias*. Traducción de Gemma Deza Guil. Barcelona: Ariel, pp. 335, $15.105 (pesos chilenos). ISBN: 978-987-3804-85-4.

DOI: https://doi.org/10.22370/rhv2019iss13pp102-108

1. Introducción

En Septiembre de 2018, la editorial Ariel publicó la traducción de Gemma Deza Guil del libro How Democracies Die, de Steven Levitsky y Daniel Ziblatt. Ambos, profesores de Harvard, han dedicado sus carreras a investigar cuestiones relacionadas con la democracia y el autoritarismo. El trabajo de Steven Levitsky se centra en Latinoamérica y el de Daniel Ziblatt en Europa desde el siglo XIX hasta el presente. En Cómo mueren las democracias, los dos autores continúan la temática pero cambian de registro geográfico. El objetivo principal del libro es investigar si los eventos que han sucedido en los últimos años han puesto en peligro la democracia en Estados Unidos. Ya que el objeto de estudio es la democracia estadounidense, podría pensarse que las consecuencias de las tesis contenidas en el libro tienen más bien un alcance local. Sin embargo, el caso de Estados Unidos es solo un ejemplo de una tendencia general que puede verse en muchas sociedades democráticas actuales: la democracia, a día de hoy, se ve amenazada por personas y partidos políticos que usan los mecanismos democráticos en contra de los mismos sistemas democráticos. En el mundo actual, las democracias fracasan "en manos no ya de generales, sino de líderes electos, de presidentes o primeros ministros que subvierten el proceso mismo que los condujo al poder." (Levitsky y Ziblatt 2018, 11).

El libro está orientado a una audiencia muy amplia, desde público general no especializado a público académico. Para aquellas personas no familiarizadas con los debates centrales en filosofía política, la obra supone una perfecta y actualizada introducción al campo. Para el público especializado es una lectura obligada ya que es una pieza fundamental del nuevo paradigma que se está consolidando los últimos años: la subversión de la democracia por los mismos mecanismos de los sistemas democráticos (ver Stanley 2017; Sunstein 2017).

2. La subversión de la democracia

Como ya se ha comentado, los autores defienden que las democracias actuales ya no caen por golpes militares, sino por los mismos medios democráticos. Es lo que los autores llaman la "subversión de la democracia": sistemas democráticos desmantelados por medios democráticos y en pos de causas democráticas. "Mientras que los dictadores de la vieja escuela solían encarcelar, enviar al exilio o incluso asesinar a sus adversarios, los autócratas contemporáneos tienden a ocultar su represión tras una apariencia de legalidad." (Levitsky y Ziblatt 2018, 101). Las democracias, en la actualidad, se desmantelan de manera paulatina, a pequeños pasos. Los autores citan varias etapas que suelen estar presentes en este desmantelamiento progresivo.

Primero, se cambian las reglas del juego de tal manera que el gobierno autocrático se asegure una permanencia continuada en el poder. Esto puede hacerse de varias maneras. Una de ellas suele ser cambiar las leyes electorales vigentes de tal manera que beneficien claramente al gobierno. Por ejemplo, el Fidesz, o Unión Cívica Húngara, de Víktor Orbán acometió una reforma de la ley electoral en 2010 que concernía tanto a los distritos como a las campañas electorales. Las consecuencias de esta nueva ley en las elecciones de 2014 fueron más que significativas. Pese a que "el porcentaje de votos del Fidesz cayó sensiblemente, de un 53 por ciento en 2010 a un 44,5 por ciento en 2014, el partido gobernante consiguió conservar su mayoría por dos tercios." (Levitsky y Ziblatt 2018, 107).

Segundo, se secuestra a aquellos organismos que tienen poder para investigar y penalizar las irregularidades del gobierno, por ejemplo, el sistema judicial, los servicios de inteligencia o las agencias tributarias. Esto puede hacerse de varias maneras, ya sea ocupando esos puestos con personas leales al partido que ostenta el poder, o simplemente comprando a las personas que ya ocupan el puesto. Si ninguna de estas dos maniobras resulta viable, siempre puede acusarse de prevaricación a aquellas personas que estén investigando al gobierno o, incluso, a aquellas que pudieran investigarlo en algún momento. Un buen ejemplo para ilustrar esto lo encontramos en la política de Fujimori. Cuando el Tribunal Constitucional peruano dictaminó que iba en contra de la Constitución que Fujimori se presentara "como candidato por tercera vez en 1997, los alidados de Fujimori en el Congreso acusaron de prevaricación a tres de los siete jueces del órgano." (Levitsky y Ziblatt 2018, 97).

Tercero, se desmantela la oposición. Esto también puede hacerse por varios medios. El modo más sencillo para hacerlo, según Levitsky y Ziblatt, es "ofrecer puestos políticos, empresariales o mediáticos destacados, favores, ventajas o, directamente, sobornos a cambio de su apoyo o, al menos, de su silencio y su neutralidad." (Levitsky y Ziblatt 2018, 99). Con aquellos políticos, empresarios o medios de comunicación que no se pueden comprar, pueden utilizarse otros medios: marginación económica y legal, demandas de difamación o campañas de desprestigio.

Cuarto, se silencia a aquellas figuras culturales (artistas, escritoras, músicos, intelectuales, deportistas) que pudieran suponer una amenaza en algún momento.

Esta nueva estrategia, una subversión silenciosa y paulatina en vez de un derrocamiento violento, implica una diferencia fundamental. Cuando la democracia se subvierte por medios democráticos, no puede identificarse un momento concreto en el que el sistema está en peligro, a diferencia de lo que sucede en los golpes militares donde el momento en el que una sociedad pierde su libertad democrática es claro. Por esta razón, es muy posible que la población de una sociedad cuya democracia está siendo subvertida no advierta que su sistema democrático está bajo amenaza. Para evitar que esto suceda, los autores proponen un test basado en cuatro indicadores clave para detectar a posibles autócratas:

1. Rechazo de las reglas democráticas del juego.
2. Negación de la legitimidad de los oponentes.
3. Tolerancia de la violencia.
4. Voluntad de restringir las libertades civiles de los opositores.

Como lo propios autores señalan, el test no es infalible. Sin embargo, sí que puede servir como mecanismo para detectar posibles futuras amenazas y, de este modo, poder actuar a tiempo para así poder poner sobre aviso a la población.

3. La subversión de la democracia en Estados Unidos

En el caso específico de Estados Unidos, el desmantelamiento de la democracia "dio comienzo hace décadas, mucho antes de que Trump descendiera por unas escaleras mecánicas para anunciar que se presentaba a la presidencia de Estados Unidos." (Levitsky y Ziblatt 2018, 170). En Estados Unidos siempre han existido autocrátas y demagogos. Los autores citan, como ejemplo, a Charles Coughlin, "un sacerdote católico antisemita cuyo programa de radio de un nacionalismo impetuoso llegaba a un público de cuarenta millones de oyentes cada semana." (Levitsky y Ziblatt 2018, 46); y a Huey Long, un gobernador y senador de Luisiana que construyó en su estado "la mayor aproximación a un estado totalitario que la república de Estados Unidos haya conocido." (Levitsky y Ziblatt 2018, 47). Tanto Coughlin como Long eran apoyados por una buena parte de la población. Sin embargo, nunca tuvieron opción real de poder llegar a postularse como candidatos. ¿Por qué? La razón era que estas figuras no contaban con prácticamente ningún apoyo dentro del partido republicano. La elección del candidato solía hacerse a puerta cerrada y, aunque pueda parecer extraño, esto tenía una ventaja: "evadía y mantenía a las figuras demostradamente inadecuadas fuera de las votaciones y de la Casa Blanca y, por ende, cumplía una función de proteger la democracia." (Levitsky y Ziblatt 2018, 51).

El momento en el que esto cambió fue el año 1968. Tras la intensificación de la guerra de Vietnam, el asesinato de Martin Luther King Jr. y de Robert F. Kennedy, y la llamada Batalla de la Avenida Michigan, el Partido Demócrata creó una comisión que elaboró una serie de recomendaciones, adoptadas por ambos partidos, que crearon un sistema de primarias presidenciales vinculantes. Eso quería decir que, "por primera vez, las personas

que elegirían a los candidatos a la presidencia de los partidos … reflejarían fielmente la voluntad de los votantes de las primarias de sus estados." (Levitsky y Ziblatt 2018, 64). Antes de eso, personas influyentes de cada partido se encargaban de apartar a personajes extremistas como Coughlin o Long en reuniones a puerta cerrada. Sin embargo, ahora esta "función de filtración" (Levitsky y Ziblatt 2018, 54) ya no existía.

Con el paso del tiempo, otros factores externos agravarían la situación antes descrita, haciendo posible que alguien como Trump, alguien que da positivo en los cuatro indicadores del test propuesto por Levitsky y Ziblatt, llegara a la Casa Blanca. Según los autores, algunas de las causas principales de la victoria de Trump fueron, entre otras, la pérdida paulatina de poder de los partidos políticos debida tanto al aumento de la financiación externa como al aumento del número de medios de comunicación alternativos; y, de manera más importante, el hecho de que la polarización de la sociedad estadounidense beneficiara a los candidatos con ideologías extremistas, como Trump.

Este último punto, la polarización política, es de suma importancia a la hora de explicar por qué los sistemas democráticos se han debilitado las últimas décadas. Como los autores señalan, "la polarización puede despedazar las normas democráticas." (Levitsky y Ziblatt 2018, 137). Varios estudios apuntan al hecho de que desde hace tiempo, especialmente en las dos últimas décadas, las sociedades de muchos países han sufrido procesos de polarización. La polarización puede entenderse de varias maneras (ver Bramson et al. 2017), pero quizá la más extendida suele ser hablar de un desplazamiento hacia los extremos del espectro político o, en otras palabras, de un "centro político en desaparición" (Abramowitz 2010). Es decir, cada día que pasa más porcentaje de la población tiende a adoptar creencias y actitudes más extremas o radicales. Las personas que simpatizan con los partidos de izquierda cada vez son más de izquierdas, y las personas que simpatizan con los partidos de derecha cada vez son más de derechas. En Estados Unidos, por ejemplo, un 49% de la población estadounidense ocupaba el centro político en 1994, pero solo un 39% en 2014. Esto ha hecho que los republicanos liberales y los demócratas conservadores, quienes ejercieron una influencia considerable en sus respectivos partidos, casi hayan desaparecido (Abramowitz 2010, 158). Esta situación conlleva un peligro importante. Como algunos estudios señalan , el aumento de la polarización puede estar asociado con un aumento de la hostilidad hacia las personas que se sitúan en el extremo político opuesto. Esto hace muy difícil que pueda haber entendimiento entre partidos de distinto signo político y que, por ejemplo, puedan llegar a acuerdos. Además, esta situación se ve agravada por otros fenómenos como las "cámaras de eco" (ver Jamieson y Capella 2010; Sunstein 2017) o las "burbujas de información" (ver Pariser 2011). Es decir, sitios, ya sean físicos o virtuales, donde la práctica totalidad de la información a la que acceden las personas que están dentro de la cámara o la burbuja es información que confirma las creencias que ya tienen previamente. Estos mecanismos de selección de información hacen que la polarización y la hostilidad hacia el oponente político aumenten.

La polarización política en Estados Unidos es uno de los casos mejor estudiados. Algunos de los hitos que los autores señalan como claves en este proceso de polarización son:

1. El triunfo de la concepción agonística de la política de Newt Gingrich en el seno del partido republicano a finales de los años 80 y principios de los 90. Poco a poco, gracias al GOPAC , una organización de entrenamiento para los candidatos republicanos, la línea dura de Gingrich fue imponiéndose en el partido. El punto culminante del cambio quizá pueda situarse en la adopción por parte del partido republicano del memorándum del GOPAC titulado "Lenguaje: Un mecanismo clave de control", en el cual se exhortaba a los candidatos republicanos a emplear "ciertos palabras peyorativas para descalificar a los demócratas, como por ejemplo: «patéticos», «enfermos», «raros», «antibandera», «antifamilia» y «traidores»." (Levitsky y Ziblatt 2018, 172).
2. El aumento del filibusterismo como práctica política habitual. Los autores señalan los dos primeros años de gobierno del demócrata Bill Clinton como el periodo donde más se ha usado este método para impedir acuerdos políticos entre los dos partidos.
3. El total abandono del respeto y la tolerancia mutua por parte de algunos sectores políticos y mediáticos del partido republicano después del 11-S. Incluso hubo quien, como Ann Coulter, llegó a vincular a los demócratas con Al-Qaeda.
4. El cuestionamiento de la legitimidad de Barack Obama como presidente en 2008 tanto por el Tea Party, como por el movimiento natalista. El natalista más conocido fue, de hecho, Donald Trump, quien ocho años más tarde se convertiría en presidente de los Estados Unidos.

La polarización política en Estados Unidos no solo atañe a cuestiones políticas, sino a cuestiones identitarias. En palabras de los autores, "los dos partidos están ahora divididos por temas de raza y religión, dos temas profundamente polarizadores que tienden a generar una mayor intolerancia y hostilidad que los temas políticos tradicionales." (Levitsky y Ziblatt 2018, 200). Ante una situación como la descrita por los autores, ¿qué se puede hacer? La respuesta a la pregunta, en su opinión, pasa por reconocer y recuperar los valores democráticos que se han perdido en este proceso: la tolerancia mutua y la contención institucional.

4. Tolerancia mutua y contención institucional: Dos valores democráticos básicos

Para los autores, una de las causas principales de que las democracias actuales se vean amenazadas está relacionada con la pérdida de dos valores democráticos básicos: la tolerancia mutua y la contención institucional. Ambas son reglas informales en el sentido de que no figuran de manera explícita ni en Constitución ni reglamento alguno. Sin embargo, las dos son fundamentales porque "todas las democracias de éxito dependen de reglas informales" (Levitsky y Ziblatt 2018, 121), siendo las dos antes comentadas las más importantes a su juicio.

La tolerancia mutua alude a la importancia que tiene el que podamos tolerar y respetar a aquellas personas que tienen ideas políticas distintas a las nuestras. Como dicen Levitsky y Ziblatt: "siempre que nuestros adversarios acaten las reglas constitucionales, aceptamos que tienen el mismo derecha a existir, competir por el poder y gobernar que nosotros." (Levitsky y Ziblatt 2018, 122). Si vemos a las personas que no comparten nuestras creencias políticas como enemigos, como personas que no merecen respeto, esto puede tener consecuencias nefastas. Primero, es probable que sea prácticamente imposible llegar a acuerdos con partidos políticos de tendencia contraria, pudiendo acabar incluso en situaciones de bloqueo total. Segundo, es posible que se adopten medidas autoritarias en un contexto en el que el oponente político es considerado una amenaza peligrosa.

La contención institucional hace referencia a "evitar realizar acciones que, si bien respetan la ley escrita, vulneran a todas luces su espíritu." (Levitsky y Ziblatt 2018, 126). Es decir, refrenarse de llevar a cabo ciertas acciones que, aunque dentro del marco de la legalidad, no se ven como adecuadas, por ejemplo porque infringen alguna ley no escrita. Los autores ponen como ejemplo claro de contención institucional el respeto por parte de los presidentes estadounidenses de emular a George Washington y no postularse como candidatos para un posible tercer mandato. Antes de 1951 la Constitución no obligaba a los presidentes de Estados Unidos a abandonar el cargo una vez que su segundo mandato hubiera concluido. El único presidente que lo hizo fue Franklin D. Roosevelt en 1940. La falta de contención a la hora de respetar la norma no escrita de no postularse como candidato por tercera vez causó gran malestar entre amplios sectores tanto demócratas como republicanos. De hecho, fue lo que causó que en 1951 se aprobara la vigésimosegunda enmienda, la cual dictamina que los presidentes tienen que abandonar la Casa Blanca tras dos mandatos.

En conclusión, la democracia en Estados Unidos, aunque amenazada en ocasiones, siempre ha sobrevivido por el respeto a "dos normas que a menudo damos por supuestas: la tolerancia mutua y la contención institucional" (Levitsky y Ziblatt 2018, 246). Aunque estos dos principios no están recogidos en ninguna ley, son fundamentales ya que posibilitan que las instituciones sigan funcionando, posibilitan que las instituciones puedan velar por el bienestar de la democracia.

5. Conclusiones

Cómo mueren las democracias es una obra de obligada lectura para entender los cambios políticos globales que han acontencido los últimos años en muchos países. Las democracias actuales han dejado de serlo en muchos casos no por la acción violenta de dictadores, sino por las acciones de presidentes y primeros ministros que usan los mismos mecanismos de los sistemas democráticos para subvertir la democracia. En sociedades polarizadas, como lo son buena parte de las sociedades actuales, la posibilidad de que autócratas demagogos lleguen al poder es aún mayor. Para combatir esto los autores de la

obra proponen una serie de criterios que permitan detectar a los autócratas antes de que lleguen al poder, así como la recuperación de dos reglas democráticas no escritas esenciales: la tolerancia mutua y la contención institucional.

Referencias bibliográficas

Abramowitz, A. I. (2010). *The Disappearing Center: Engaged Citizens, Polarization, & American Democracy*. New Haven and Londres: Yale University Press.

Bramson, A. et al. (2017). Understanding Polarization: Meanings, Measures, and Model Evaluation. *Philosophy of Science*, 84(1): 115-159.

Jamieson, K. H., Cappella, J. N. (2010). *Echo Chamber: Rush Limbaugh and the Conservative Media Establishment*. Nueva York: Oxford University Press.

Levitsky, S., Ziblatt, D. (2018). *Cómo Mueren las Democracias*. Barcelona: Ariel.

Pariser, E. (2011). *The Filter Bubble: What Internet is Hiding from You*. Londres: The Penguin Books.

Stanley, J. (2017). *How Propaganda Works*. Nueva Jersey: Princeton University Press.

Sunstein, C. R. (2017). *#Republic: Divided Democracy in the Age of Social Media*. New Jersey: Princeton University Press.

David Bordonaba Plou

FONDECYT / Universidad de Valparaíso
david.bordonaba@uv.cl