UNICAMP

MARCOS ANTONIO ALVES

LÓGICA E INFORMAÇÃO

uma análise da consequência lógica a partir de uma perspectiva quantitativa da informação

CAMPINAS
2012

UNICAMP

Universidade Estadual de Campinas
Instituto de Filosofia e Ciências Humanas
Programa de Doutorado em Filosofia

MARCOS ANTONIO ALVES

LÓGICA E INFORMAÇÃO

uma análise da consequência lógica a partir de uma perspectiva quantitativa da informação

ITALA MARIA LOFFREDO D'OTTAVIANO

Tese de Doutorado apresentada ao Instituto de Filosofia e Ciências Humanas, para obtenção do Título de Doutor em Filosofia

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DA TESE DEFENDIDA PELO ALUNO MARCOS ANTONIO ALVES E ORIENTADA PELA PROFA. DRA ÍTALA MARIA LOFFREDO D'OTTAVIANO.

CPG, _____/_____/______

CAMPINAS
2012

FICHA CATALOGRÁFICA ELABORADA POR
MARIA JÚLIA MILANI RODRIGUES – CRB8/2116
BIBLIOTECA DO IFCH UNICAMP

Alves, Marcos Antonio, 1975-
AL87L Lógica e informação : uma análise da consequência lógica a partir de uma perspectiva quantitativa da informação / Marcos Antono Alves. - - Campinas, SP : [s. n.], 2012.

Orientadora: Ítala Maria Loffredo D'Ottaviano
Tese (doutorado) - Universidade Estadual de Campinas, Instituto de Filosofia e Ciências Humanas.

1. Filosofia. 2. Lógica. 3. Epistemologia. 4. Lógica simbólica e matemática. 5. Teoria da informação. I. D'Ottaviano, Itala Maria Loffredo, 1944-. II. Universidade Estadual de Campinas. Instituto de Filosofia e Ciências Humanas. III. Título.

<u>Informação para Biblioteca Digital</u>

Título em Inglês: Logic and information: an approach quantitative Informational of logical consequence
Palavras-chave em inglês:
Philosophy
Logic
Epistemology
Logic, Symbolic and mathematical
Information theory
Área de concentração: Filosofia
Titulação: Doutor em Filosofia
Banca examinadora:
Ítala Maria Loffredo D'Ottaviano [Orientador]
Maria Eunice Quilici Gonzalez
Lauro Frederico Barbosa da Silveira
Hércules de Araújo Feitosa
Fábio Maia Bertato
Data da defesa: 14/12/2012
Programa de Pós-Graduação: Filosofia

UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE FILOSOFIA E CIÊNCIAS HUMANAS

A Comissão Julgadora dos trabalhos de Defesa de Tese de Doutorado, em sessão pública realizada em 14 de dezembro de 2012, considerou o candidato MARCOS ANTONIO ALVES aprovado.

Este exemplar corresponde à redação final da Tese defendida e aprovada pela Comissão Julgadora.

Profa. Dra. Itala Maria Loffredo D'Otaviano

Profa. Dra. Maria Eunice Quilici Gonzalez

Prof. Dr. Lauro Frederico Barbosa da Silveira

Prof. Dr. Hercules de Araújo Feitosa

Prof. Dr. Fábio Maia Bertato

Dedico este trabalho a meu amado filho
João Pedro

Agradecimentos

A realização deste trabalho deve-se à contribuição de diversas pessoas, em diferentes aspectos. Desejo agradecer em especial:

à minha querida orientadora, Ítala Maria Loffredo D'Ottaviano, sem a qual este trabalho não chegaria ao ponto em que está. Agradeço pela paciência, força, compreensão, carinho e afeto dedicados ao longo deste trabalho. É um imenso orgulho e prazer ter sido orientado por tal exemplo de dedicação, competência e profissionalismo acadêmicos;

aos membros da banca examinadora, cujas valorosas contribuições propiciaram que este trabalho contenha menos imperfeições;

aos colegas e docentes do CLE, pelo convívio e contribuição na minha formação acadêmica;

à professora Maria Eunice Quilici Gonzalez, querida mestra cujo apoio acadêmico, profissional e pessoal são fundamentais já há longo tempo em minha vida.

à querida Edna Alves de Souza, pelos anos de convivência e parceria;

aos amigos Carlos Sergio Bastianik Fernandes e Elaine Carvalho Fernandes pela presença nos momentos difíceis;

a todos aqueles que, direta ou indiretamente, contribuíram para que este trabalho fosse realizado;

por fim, e mais fundamental, a meu amado filho João Pedro de Souza Alves. A sua existência é altamente informativa; diferença que faz muita diferença.

Obrigado a todos!!!

"... postas algumas coisas,
outras se seguem necessariamente..."
(Aristóteles)

"... mas para os meus olhos experimentados
cada marca na sua superfície
tinha um significado especial, disse Holmes."
(Conan Doyle)

Resumo

Nosso objetivo nesta Tese é desenvolver uma definição da consequência lógica a partir de uma perspectiva quantitativa da informação. O trabalho pode ser dividido em duas partes. Na primeira, que consiste dos três capítulos iniciais, fazemos um estudo crítico de algumas das principais concepções usuais de consequência lógica. No primeiro capítulo expomos três das características centrais da consequência lógica, quais sejam, necessidade, formalidade e anterioridade. Apresentamos uma noção geral de consequência lógica, a partir da qual classificamos as diferentes noções de consequência lógica em clássicas e não-clássicas. Nos dois próximos capítulos tratamos da consequência lógica a partir das perspectivas sintática e semântica, analisando em que medida elas satisfazem as três características acima enunciadas. Na segunda parte, constituída dos três últimos capítulos, desenvolvemos a nossa proposta. No quarto capítulo expomos criticamente a concepção de informação a ser utilizada na Tese. No quinto capítulo construímos uma semântica probabilística para a lógica sentencial clássica, mostrando os seus principais resultados. A partir desta semântica, definimos, no sexto capítulo, a quantidade de informação em uma fórmula da lógica sentencial clássica e a consequência lógica probabilística. Feito isso, definimos a consequência lógica informacional, demonstrando os seus principais resultados.

Palavras-chave: Consequência lógica, verdade, probabilidade, informação, consequência lógica informacional.

Abstract

Our goal in this work is to develop a definition of logical consequence from a quantitative perspective of information. The work can be divided into two parts. The first, consisting of three chapters, we make a critical study of some key concepts usual logical consequence. In the first chapter we expose three of the central features of logical consequence, namely, necessity, formality and apriority. We present a general notion of logical consequence, from which classify the different notions of logical consequence in classical and non-classical. In the next two chapters deal with the logical consequence from the syntactic and semantic perspectives, analyzing the extent to which they meet the three above features. In the second part, which consists of the last three chapters, we developed our proposal. In the fourth chapter critically expose the concept of information to be used in the thesis. In the fifth chapter we build a probabilistic semantics for classical sentential logic, showing its main results. From this semantics, we define, in the sixth chapter, the amount of information in a formula of classical sentence logic and probabilistic logical consequence. That done, we define the informational logical consequence, showing its main results.

Keywords: logical consequence, truth, probability, information, informational logical consequence.

Sumário

Introdução

Pode-se dizer que a Lógica, enquanto disciplina, tenha sido criada por Aristóteles. Foi o Estagirita quem, pela primeira vez, construiu um sistema razoável para tratar de argumentos válidos. Seu objetivo, especialmente nos *Primeiros Analíticos*, era mostrar que coisas se seguem necessariamente de outras coisas. Em suma, ele propunha uma noção de consequência lógica, de argumento válido, estabelecendo as suas principais características.

Até o século XIX poucas foram as alterações na Lógica. No final daquele século começaram a surgir novos e revolucionários métodos para avaliação da validade de argumentos, ou da satisfação da consequência lógica entre um dado conjunto de sentenças e uma sentença. Foram construídas linguagens e ferramentas capazes de tratar de uma gama muito mais ampla de argumentos do que a Silogística Aristotélica.

No decorrer do século XX as discussões sobre as características e propriedades da consequência lógica se intensificaram. Surgiram diferentes noções de consequência lógica, seja em relação às suas propriedades, seja em relação àquilo que se considera um argumento válido.

Apesar das diferenças, em termos semânticos, ao longo dos tempos, a consequência lógica foi costumeiramente definida a partir da manutenção da verdade das premissas para a conclusão. Nosso objetivo central nesta tese consiste no desenvolvimento de uma análise da consequência lógica a partir de uma perspectiva quantitativa da informação.

Nosso trabalho, cujo fio condutor é a noção de consequência lógica, pode ser dividido em duas partes. Na primeira delas, que consiste dos três capítulos iniciais, fazemos um estudo crítico de algumas das principais concepções usuais de consequência lógica. Na segunda parte, constituída dos três últimos capítulos, desenvolvemos a nossa proposta.

No primeiro capítulo exibimos, inicialmente, aquelas que foram, ao longo da história, consideradas características elementares da consequência lógica: a necessidade, a formalidade e a anterioridade. Em seguida expomos um conceito geral de consequência lógica, baseados na proposta de Tarski. Como é bem conhecido, de acordo com essa concepção, para uma consequência lógica ser considerada clássica, ou Tarskiana, precisa satisfazer, dentre outras, três propriedades: reflexividade, monotonicidade e transitividade.

No segundo capítulo expomos uma definição de consequência lógica sintática, a fim de averiguar se ela satisfaz as características da consequência lógica expostas no capítulo anterior. Iniciamos apresentando a ideia de um sistema formal lógico e estabelecemos uma classificação destes sistemas em clássicos e não-clássicos. Por fim, expomos uma definição de consequência lógica sintática e analisamos as suas principais características.

No terceiro capítulo tratamos da consequência lógica semântica. Se, por um lado, na parte sintática, a consequência lógica é definida unicamente através da manipulação de símbolos a partir de regras de inferência, na parte semântica ela é definida a partir da "interpretação" das expressões da linguagem de um sistema formal lógico. Neste capítulo apresentamos algumas perspectivas semânticas da consequência lógica, averiguando em que medida elas satisfazem as características da necessidade, formalidade e anterioridade.

No quarto capítulo tratamos da noção quantitativa da informação, tal como definida por pesquisadores como Shannon e adotada na Teoria Matemática da Comunicação. Começamos pela exposição de um modelo de comunicação e pela definição de uma fonte de informação, tomada como referência para a noção de situação, a ser definida posteriormente. Depois de expormos a noção de informação subjacente à proposta de Shannon e de estabelecermos uma definição da medida informacional de uma fonte, terminamos o capítulo com uma análise crítica da noção quantitativa da informação.

No quinto capítulo introduzimos uma semântica para a lógica sentencial clássica baseada na noção de valor probabilidade. Apresentamos, inicialmente, alguns conceitos básicos de uma teoria axiomática usual de probabilidades. Em seguida propomos uma "interpretação" das fórmulas da linguagem da lógica sentencial clássica, relacionando-as a eventos de um experimento aleatório e introduzimos a definição de valor probabilidade de uma fórmula. Ao longo do capítulo mostramos uma série de resultados oriundos desta proposta.

Por fim, no sexto capítulo, desenvolvemos uma perspectiva da consequência lógica baseada na noção de quantidade de informação de um dado conjunto de premissas e de uma conclusão, que depende da noção de valor probabilidade. Definimos o valor informacional de uma fórmula, além das definições de consequência lógica probabilística e consequência lógica informacional. Para cada uma delas, provamos alguns resultados que consideramos relevantes para os objetivos do nosso trabalho.

Capítulo 1

Consequência lógica

1 Apresentação

Neste capítulo expomos uma visão geral da noção de consequência lógica. Iniciamos, na próxima seção, apresentando três características comumente consideradas fundamentais da consequência lógica: a necessidade, a formalidade e a anterioridade. Na terceira seção apresentamos uma definição clássica de consequência lógica, a ser tomada como ponto de partida para a classificação das distintas definições de consequência lógica. Tal definição, sintetizada por Tarski, caracteriza-se por apresentar certas propriedades como a reflexividade, a monotonicidade e a transitividade. Em seguida, tratamos das características dos elementos envolvidos em uma relação de consequência lógica, analisando três possíveis candidatos a satisfazê-las: a sentença, a proposição e o enunciado, tomando como ponto de partida a noção de valor de verdade. Devido aos nossos interesses neste trabalho, assumiremos que os elementos de uma consequência lógica são as fórmulas de uma linguagem. Por isso, na última seção, apresentamos uma concepção de sistema axiomático, expondo, genericamente, os elementos que o compõem. Eles servirão de ponto de partida para as definições de consequência lógica apresentadas nos próximos capítulos.

2 Características pré-teóricas da consequência lógica

A consequência lógica pode ser pensada de diferentes modos. Ela pode ser definida, por exemplo, como um operador sobre um conjunto não vazio ou como algum tipo de relação entre elementos de um conjunto não vazio.

Em termos intuitivos, a consequência lógica é frequentemente concebida como uma relação que se mantém entre um dado conjunto de orações de uma linguagem e uma oração da mesma linguagem. Na Lógica, é comum defini-la de modo semelhante ao intuitivo, como uma relação entre um dado conjunto, possivelmente vazio ou mesmo infinito, de sentenças, ou de fórmulas, de uma linguagem, geralmente formal, e uma sentença da mesma linguagem. As orações ou as sentenças pertencentes ao dado conjunto são denominadas *premissas* e a outra oração ou sentença é denominada *conclusão*. Nessa relação, as premissas possuem o papel de fundamentar, sustentar, apoiar a conclusão.

Neste trabalho consideraremos esta perspectiva sentencial da Lógica sobre a consequência lógica, considerando a visão intuitiva como comparação ou exemplificação na análise do objeto de nossa investigação.

Quando reunidas adequadamente em uma sequência, as premissas e a conclusão constituem um *argumento*. Diremos que uma conclusão é consequência lógica de um conjunto de premissas se, e somente se, o argumento composto pela sua reunião é logicamente válido ou, simplesmente, válido.

Nesta seção apresentamos três características consideradas essenciais da consequência lógica por diversos pensadores ao longo da história. Elas são a necessidade, a formalidade e a anterioridade ou *aprioridade*.

2.1 Necessidade

A característica da *necessidade* foi descrita por diversos pensadores, como Aristóteles (1984, Primeiros Analíticos, 24b 19-22; Tópicos 100a, 25-6), Leibniz (1989, I. 28 1989, p. 209) e Tarski (1956d, p. 414-5). De acordo com Etchemendy (1999, p. 81),

> A mais importante característica da consequência lógica, como nós a entendemos ordinariamente, é uma relação modal que se mantém entre as sentenças que implicam e a implicada. As premissas de um argumento logicamente válido não podem ser verdadeiras se a conclusão for falsa; dizemos que tais conclusões "seguem-se necessariamente" de suas premissas.

A necessidade indica o caráter de confiabilidade, de garantia absoluta da conclusão a partir das premissas. Para que uma sentença seja consequência lógica de um dado conjunto de sentenças, a relação ou a inferência deve ser mantida em qualquer circunstância logicamente possível.

Uma *circunstância* pode ser uma valoração, estrutura, interpretação de certas expressões de uma linguagem, como feito usualmente na Lógica, um mundo possível, uma realidade, um conjunto de elementos ou de fatos no mundo, uma situação, como definiremos no quinto capítulo. Uma circunstância *logicamente possível* respeita certas leis ou princípios lógicos pré-estabelecidos. Na lógica clássica aristotélica, por exemplo, os principais princípios a serem satisfeitos são os da identidade, do terceiro excluído e da não-contradição. Se uma circunstância contrariar algum deles, não é logicamente possível, segundo esta lógica.

A característica da necessidade da consequência lógica indica a sua independência de qualquer circunstância, mantendo-se inalterada tanto naquilo que ocorre quanto naquilo que pode ocorrer. Ela não deve ser estabelecida, por exemplo, com base em mundos específicos. Para “Érico Veríssimo é brasileiro” ser consequência lógica do conjunto unitário constituído pela oração “Érico Veríssimo é gaúcho”, por exemplo, toda circunstância logicamente possível deveria garantir a validade da inferência ou a satisfação da relação de consequência lógica entre estas duas orações. Quando a relação é definida, por exemplo, em termos semânticos, pela manutenção da verdade da conclusão a partir da suposição da verdade das premissas, como usualmente feito por Mendelson (1964) ou Shoenfield (1967), pressupondo a noção de verdade como correspondência, a oração “Érico Veríssimo é brasileiro” deveria ser verdadeira sempre que “Érico Veríssimo é gaúcho” fosse verdadeira.

No exemplo específico acima pode-se observar a manutenção da verdade no mundo *real*. No entanto, existem circunstâncias logicamente possíveis em que a premissa é verdadeira e a conclusão é falsa. Este seria o caso do mundo em que o Rio Grande do Sul não fizesse parte do território do Brasil. Érico Veríssimo continuaria sendo gaúcho, mas não seria brasileiro.

A mudança de circunstâncias geralmente implica uma mudança de significado dos termos ou da natureza dos objetos referidos por eles. A modificação do mundo *real* no último exemplo acima leva à modificação da referência ou da extensão de “brasileiro”. Indiretamente, os termos “Érico Veríssimo” e “gaúcho” também tiveram seu significado modificado, mesmo permanecendo inalterados em sua estrutura simbólica.

A necessidade não pode ser analisada sem algum tipo de modificação de aspectos semânticos da sentença, ao interpretá-la de acordo com a variação das circunstâncias. Se fosse proibida qualquer modificação deste tipo, o único mundo a ser analisado seria o mundo *real*, o que tornaria o caráter de necessidade contraditório: exigiria a consideração de todo mundo possível e proibiria a análise de qualquer mundo distinto do *real*.

A consideração de circunstâncias distintas, tal como exigida pela característica da necessidade, pode produzir mudanças semânticas nas sentenças, resultando em modificações do significado ou do conteúdo dos seus termos.

A segunda característica costumeiramente atribuída à consequência lógica explicita justamente que ela não deve ser definida a partir do conteúdo das sentenças ou orações, mas sim da sua forma.

2.2 Formalidade

A *formalidade* explicita a ideia de que a consequência lógica é identificada ou definida a partir da estrutura ou da forma dos seus elementos constituintes. Corcoran (2009, p. 4) atribui a Aristóteles a descoberta dessa característica.

A formalidade, explica Field (1989, p. 116-9, 239), exige a análise de toda compreensão logicamente possível das sentenças envolvidas. Uma *compreensão logicamente possível* de um conjunto de sentenças pode ser entendida como uma série de operações, do tipo substituições ou interpretações dos termos das sentenças, que não modificam a sua forma. Para manter a forma inalterada, termos iguais seriam interpretados ou substituídos por termos iguais e termos diferentes por termos diferentes.

A consequência lógica se caracteriza por abstrair o conteúdo dos seus elementos. Dependendo do ponto de vista, desconsidera qualquer conteúdo ou considera todo conteúdo. Desconsidera porque a consequência lógica deve se basear na distribuição dos termos, independente do seu conteúdo, exceto talvez de alguns termos especiais. Por outro lado, explicita Read (1994, p. 49), para uma sentença com conteúdo específico ser consequência lógica de um conjunto específico de premissas, qualquer sentença que tiver a mesma forma, independente do seu conteúdo, também deve ser consequência lógica de todo conjunto de premissas que partilham da mesma forma desse conjunto específico. Ou seja, a consequência lógica se aplica a qualquer conteúdo.

A *forma* de uma sentença pode ser definida a partir da sua estrutura e da relação entre os seus termos. De acordo com Tarski (1956c, p. 157), a descrição da estrutura de uma sentença de uma linguagem (ou de qualquer outra expressão) consiste na explicitação das expressões que a constituem, além dos sinais que compõem cada expressão e da ordem em que se apresentam. Assim, "Érico Veríssimo é Érico Veríssimo" é uma sequência composta de cinco palavras, todas elas, exceto a terceira, são substantivos, de modo que a primeira palavra é idêntica à penúltima e a segunda é idêntica à última e, por fim, a primeira palavra é composta pelos símbolos "É", "r", "i", "c", "o" (nesta ordem), a segunda é composta pelos símbolos “V”, “e”, etc. e assim sucessivamente até a última palavra da sentença em questão.

Para Cohen & Nagel (1949, p. 11),

> Uma forma é, em geral, algo no qual um número de diferentes objetos ou operações concorda (embora difiram em outros aspectos), de modo que os objetos podem ser variados e ainda assim a forma permanecer a mesma. A implicação lógica [consequência lógica] é formal no sentido que ela se aplica a todas as proposições [sentenças], não importa o quanto difiram, desde que mantenham entre si certas relações.

A diferença entre as sentenças explicitadas na citação acima se refere ao seu conteúdo, àquilo que elas tratam, dizem ou se referem. As relações a serem mantidas são, em geral, determinadas pela distribuição dos símbolos lógicos e pela natureza dos símbolos não-lógicos.

A forma de uma expressão pode ser pensada como propriedade inerente a ela, possuindo um caráter absoluto, ou como uma propriedade relativa a escolhas determinadas. Uma perspectiva *absolutista* da forma lógica pressupõe a existência de propriedades substanciais às sentenças, definidoras da sua estrutura, que a estabelecem de modo universal e necessário. Uma sentença teria sempre a mesma forma, independentemente de quaisquer variações, inclusive linguísticas.

Uma perspectiva *relativista* da forma lógica a torna dependente de variações, especialmente as linguísticas. Como a estrutura de uma sentença é determinada com base nas relações entre os seus termos, pode ocorrer desses termos apresentarem propriedades distintas em diferentes linguagens. Logo, é possível também que as relações entre os termos variem de uma linguagem para outra, tornando a forma dependente da linguagem subjacente.

Se a forma for relativa à linguagem, deve haver meios de defini-la em uma dada linguagem específica. Um meio de apresentar essa definição seria através da classificação dos termos em categorias, abstraindo a forma a partir dessa classificação. Este seria um trabalho muito dispendioso. As linguagens consideradas mais expressivas geralmente possuem uma quantidade infinita de termos de natureza muito distinta, sem considerar que, em certas linguagens, como as naturais, eles podem ser ambíguos, imprecisos, indefinidos, dado o caráter dinâmico destas linguagens.

Outro modo, muito comum, de definir a forma seria através da escolha de expressões referenciais, chamadas *termos lógicos* ou *símbolos lógicos*, dependendo da perspectiva, se intuitiva ou lógica, por exemplo, ou da própria natureza destas expressões, que podem ser palavras ou determinados símbolos isolados. Neste sentido, duas expressões quaisquer de uma linguagem compartilham a mesma forma se possuírem os mesmos termos ou símbolos lógicos, cada um deles relacionado do mesmo modo em ambas as expressões com os demais elementos da expressão.

Uma das dificuldades desta proposta consiste na identificação dos termos lógicos. De acordo com Hanson (1997, p. 377), pode-se considerar que tais termos possuem características que os tornam eminentemente lógicos ou que são escolhidos arbitrariamente, dependendo, em grande parte, dos objetivos almejados em uma pesquisa lógica.

Nós seguiremos a concepção mais comum de forma de uma expressão, definindo-a a partir da sua estrutura, da categoria de suas expressões e da relação entre elas, com base na escolha dos símbolos lógicos de sua linguagem subjacente.

A título de exemplo, consideremos uma linguagem da lógica sentencial clássica usual. As fórmulas “$A_1 \wedge A_2$” e “$A_1 \rightarrow A_2$” não partilham da mesma forma lógica. Ainda que elas possuam as mesmas fórmulas atômicas, tratem do mesmo conteúdo em uma mesma circunstância, poderíamos dizer, estão conectadas por símbolos lógicos diferentes. Elas são estruturalmente distintas. Enquanto a primeira é uma conjunção, a segunda é uma implicação material, conforme a concepção usual.

Já as fórmulas “$A_1 \rightarrow A_2$” e “$A_3 \rightarrow A_4$” partilham da mesma forma lógica. São constituídas por duas fórmulas atômicas distintas unidas pelo conectivo da implicação material. Este exemplo ilustra a possibilidade de sentenças apresentarem conteúdo distinto (tratar de matérias diferentes) e compartilharem da mesma forma lógica. Mostra também que elas podem apresentar estruturas distintas, no sentido de conter símbolos diferentes, pertencentes à mesma categoria, e partilhar da mesma forma.

O caso do último exemplo acima não se aplica completamente a “$A_1 \rightarrow A_2$” e “$A_1 \rightarrow A_1$”. A rigor, essas duas sentenças não partilham das mesmas fórmulas, dado que uma conecta duas fórmulas possivelmente distintas, enquanto a segunda conecta uma fórmula a si mesma. Elas são estruturalmente distintas, embora ambas as sentenças sejam implicativas. Poderíamos dizer, neste caso, que ambas as sentenças partilham da mesma forma num sentido estendido do termo, mas não num sentido estrito.

A observação acima também se aplica a fórmulas como "$(A_1 \wedge A_2) \rightarrow A_2$" e "$A_1 \rightarrow (A_2 \wedge A_2)$" que, em outra notação, como a polonesa, podem ser escritas, respectivamente, como "$\rightarrow\wedge A_1A_2A_2$" e "$\rightarrow A_1\wedge A_2A_2$", embora apresentem os mesmos símbolos, não possuem a mesma forma no sentido estrito do termo, dada a diferença na distribuição deles. A relação entre os símbolos não é a mesma nas duas sentenças. Elas partilham da mesma forma no âmbito macroscópico, mas não no microscópico. Ambas possuem a forma implicativa, mas o antecedente da primeira é uma sentença conjuntiva o da segunda é uma sentença atômica. Essa diferença na forma pode produzir resultados muito diferentes em relação às características das sentenças, ou mesmo em relação à satisfação da consequência lógica entre sentenças. No caso das duas sentenças do exemplo em questão, tem-se que, de acordo com as definições semânticas usuais da lógica sentencial clássica, a primeira é uma tautologia, enquanto a segunda é uma sentença contingente; ainda ocorre que, nesta lógica, a primeira sentença é consequência lógica, seja em termos semânticos ou sintáticos, de qualquer conjunto de premissas, o que não ocorre com a segunda sentença.

A alteração do conjunto de símbolos lógicos de uma linguagem pode modificar a forma das suas expressões. Imaginemos uma linguagem de uma lógica sentencial clássica em que apenas as fórmulas atômicas "A_1" e "A_2" fossem os símbolos lógicos e os conectivos "$\wedge$" e "$\rightarrow$" pertencessem à mesma categoria, de símbolos não-lógicos. Então as sentenças "$A_1 \wedge A_2$" e "$A_1 \rightarrow A_2$" compartilhariam a mesma forma, enquanto "$A_1 \rightarrow A_2$" e "$A_2 \rightarrow A_1$" passariam a apresentar formas distintas, o que parece contrariar a intuição.

É possível sentenças escritas em linguagens ou notações distintas partilharem da mesma forma. Esse é o caso de "$A_1 \wedge A_2$", conforme notações usuais atuais, e "&PQ", conforme as primeiras notações da lógica sentencial clássica, pressupondo uma linguagem na qual o símbolo "&" é o símbolo para a conjunção e "P" e "Q" são sentenças atômicas conectadas por "&".

Para Hanson (1997, p. 376) "... as formas de argumentos são sempre os representantes gerais de classes de argumentos individuais." Ao tratar da forma de um argumento específico, provando-lhe uma dada propriedade, por exemplo, que ele é válido, provamos tal propriedade para todos aqueles que possuem a mesma forma.

A necessidade e a formalidade pressupõem que a consequência lógica deve valer para cada circunstância e para cada compreensão logicamente possível. A terceira característica, que passamos a expor a seguir, explicita que ela deve também ser reconhecida aprioristicamente.

2.3 Anterioridade

A *anterioridade* ou *aprioridade* pressupõe a independência da consequência lógica de conhecimentos individuais ou prévios, do grau informacional sobre o mundo ou significado dos termos, de relações de conteúdo ou algum tipo de experiência referente aos fatos expressos pelas premissas e conclusão. De acordo com Corcoran (2009, p. 4), ela é independente do conteúdo no sentido de que não é preciso conhecimento do assunto tratado pelos seus elementos.

A consequência lógica deve ser determinada *a priori*, antes de qualquer conhecimento sobre as relações entre os termos das premissas e conclusão. Deve ser descoberta independentemente de qualquer experiência, diz Hanson (1997, p. 377). Não pode depender do conteúdo factual das sentenças ou dos estados cognitivos de alguém a seu respeito, enuncia Mates (1968, p. 03).

Segundo Etchemendy (1999, p. 89)

> Poderíamos afirmar, não implausivelmente, que é somente devido à relação *a priori* entre as premissas e a conclusão de um argumento válido que julgamos a última seguir-se necessariamente das primeiras e, por isso, consideramos o argumento válido. Sob essa perspectiva, uma consequência necessária que não pudesse ser reconhecida como tal *a priori* jamais poderia ser qualificada como consequência lógica.

A anterioridade poderia ser pensada como uma característica derivada das outras duas. Isto porque, se a consequência lógica fosse reconhecida *a posteriori*, a sua definição dependeria da análise prévia de algum mundo ou compreensão particular. Isso contradiria a característica da necessidade ou da formalidade. Se a consequência lógica pressupõe a manutenção da verdade em todo mundo e compreensão logicamente possíveis, não pode depender da análise de um mundo ou compreensão particular, ou seja, não pode ser reconhecida *a posteriori*. Logo, satisfaz a anterioridade.

Em nossa opinião o argumento acima é falacioso. Nem sempre é possível inferir a anterioridade a partir da dependência de uma totalidade de casos ou fatos. Consideremos a sentença “Todo cisne é branco”. Embora a verdade desta sentença dependa da cor de cada cisne, não podemos inferir que a sua verdade possa ser determinada sem a observação ou análise prévia de cada cisne. Do mesmo modo, do fato da consequência lógica depender da manutenção da verdade em toda circunstância e compreensão logicamente possíveis não podemos inferir que seu reconhecimento seja *a priori*. Poderia ser exigida a análise prévia da manutenção da verdade em cada circunstância ou compreensão, como no caso dos cisnes, o que resultaria em um reconhecimento *a posteriori* da consequência lógica. Por esse motivo, consideramos a *aprioridade* uma característica independente das outras duas.

As três características da consequência lógica examinadas nesta seção podem ser percebidas na seguinte citação de Tarski (p. 414-5):

> Considere uma classe Γ qualquer de sentenças e uma sentença φ que se segue das sentenças desta classe. A partir de um ponto de vista intuitivo jamais pode ocorrer que Γ seja constituído apenas de sentenças verdadeiras e φ seja falsa. Além disso, uma vez que estamos interessados aqui com o conceito de consequência lógica, isto é *formal* e, assim, com uma relação que deve ser unicamente determinada pela forma das sentenças entre as quais ela se mantém, esta relação não pode ser influenciada de nenhum modo pelo conhecimento empírico e, em particular, pelo conhecimento dos objetos para os quais a sentença φ ou as sentenças de Γ se referem. A relação de consequência não pode ser afetada ao substituirmos as designações dos objetos referidos nestas sentenças pelas designações de outros objetos.

Essas três características também costumam ser atribuídas aos teoremas e às sentenças logicamente válidas de uma teoria ou de um sistema formal lógico, a serem definidas nos próximos capítulos. O caráter de validade de uma fórmula deve ser identificado a partir da sua forma, independente de circunstância ou de conhecimentos prévios.

As três características enunciadas acima costumam ser adotadas como parâmetros a serem seguidos por qualquer definição razoável de consequência lógica. Pode-se construir ou propor diferentes definições de consequência lógica razoáveis, dependendo dos objetivos teóricos ou das noções subjacentes a serem capturadas. Assim, uma definição pode ser ou não ser monotônica, aceitar apenas conjuntos finitos de premissas ou também considerar conjuntos infinitos. Ela pode ser definida a partir de diferentes âmbitos, seja formal ou informal, a partir de uma perspectiva sintática, semântica ou pragmática. Apesar das diferenças estabelecidas por diferentes definições, para ser aceita como razoável, deve satisfazer as características da necessidade, formalidade e anterioridade.

A necessidade, formalidade e anterioridade serão tomadas como base para a análise das diferentes definições ou perspectivas da consequência lógica examinadas neste trabalho. Na próxima seção apresentamos uma definição geral de consequência lógica, a partir da qual classificaremos as diferentes definições de consequência lógica em clássicas e não-clássicas.

3 Um conceito geral de consequência lógica

Dentre os pesquisadores contemporâneos a procurar estabelecer as propriedades gerais da consequência lógica, a proposta de Tarski (1956) acabou se tornando a mais conhecida e aceita na comunidade científica. Em sua ideia inicial, em artigos escritos originalmente em 1930, Tarski (1956b, p. 63-4; 1956a, p. 31) expressa axiomaticamente as propriedades que ele supõe elementares da noção de consequência lógica para um conjunto de sentenças de uma linguagem. Ele estabelece, por exemplo, que o conjunto de sentenças analisadas deve ser enumerável e assume axiomaticamente a compacidade, ou seja, embora uma sentença possa ser consequência de um conjunto infinito de sentenças, ela deve ser consequência lógica de um subconjunto finito de fórmulas do conjunto original.

O conjunto de propriedades básicas da consequência lógica foi posteriormente modificado pelo próprio Tarski (1956c; 1956d), em textos escritos em 1933 e 1936. Ele elimina desse conjunto tanto a enumerabilidade quanto a compacidade, podendo derivá-la do novo conjunto de propriedades básicas.

Apesar de Tarski se referir a sentenças de uma linguagem, apresentamos abaixo a definição de consequência lógica para um dado conjunto não vazio qualquer de elementos, como o fazem, por exemplo, Feitosa & D’Ottaviano (2004, p. 71).

Definição 1.3.1 (Relação de consequência lógica Tarskiana)

Seja $\Im$ um conjunto não vazio. Então uma *relação de consequência Tarskiana sobre* $\Im$ é um conjunto R de pares ordenados tal que $R \subseteq \wp(\Im) \times \Im$, que satisfaz as seguintes propriedades, para todo $\Gamma, \Delta \subseteq \Im$ e para todo $x, y \in \Im$:

(T1) Se $x \in \Gamma$, então $\langle \Gamma, x \rangle \in R$.

(T2) Se $\Gamma \subseteq \Delta$ e $\langle \Gamma, x \rangle \in R$, então $\langle \Delta, x \rangle \in R$.

(T3) Se $\langle \Delta, y \rangle \in R$, para cada $y \in \Gamma$, e $\langle \Gamma, x \rangle \in R$, então $\langle \Delta, x \rangle \in R$.

A expressão "$\langle \Gamma, x \rangle \in R$" é lida como "x é consequência lógica de Γ", em que x é a conclusão e os elementos de Γ são as premissas.

A partir de (T1) – (T3) é possível provar outras propriedades da relação de consequência Tarskiana, como faz Tarski (1956) e como podemos encontrar em Feitosa (1998). De (T1) e (T3), por exemplo, provamos a propriedade da *idempotência*, segundo a qual um elementos de $\mathfrak{I}$ é consequência lógica de um subconjunto de $\mathfrak{I}$ se, e somente se, for consequência lógica do conjunto das fórmulas que são consequência lógica deste subconjunto.

A propriedade (T1), denominada de *reflexividade*, estabelece que qualquer elemento de um conjunto é consequência deste conjunto. Em particular, um elemento é sempre consequência lógica de si mesmo. (T2), chamada *monotocidade*, expressa que se um elemento é consequência lógica de um dado conjunto de elementos, então a adição de mais elementos nesse conjunto ainda mantém a relação de consequência lógica com o elemento anterior. O acréscimo de novas hipóteses, dispensáveis para a consequência lógica, neste caso, mantém a relação de consequência entre o novo conjunto de hipóteses e o elemento x. (T3), denominada *transitividade*, expressa que as consequências das consequências de um dado conjunto também são consequência direta dele.

Enquanto relação, a consequência lógica entre elementos de um dado conjunto $\mathfrak{I}$ pode ser entendida como um conjunto de pares ordenados cujo primeiro elemento de cada par é um subconjunto de elementos de $\mathfrak{I}$ e o segundo é um elemento específico de $\mathfrak{I}$.

Considerando a linguagem de uma lógica quantificacional clássica usual como exemplo de uma relação de consequência não-Tarskiana sobre o seu conjunto $\mathcal{I}$ de fórmulas podemos citar: $\langle Aa, \forall x\, Ax \rangle \in R$; $\langle \forall x\, Ax, Ab \rangle \in R$, mas $\langle Aa, Ab \rangle \notin R$. Este seria o caso de uma regra do tipo "Introdução do Quantificador Universal", entendida como uma espécie de "Inferência Indutiva", em que, de um caso particular, é permitido inferir uma fórmula universal. No entanto, não há uma regra que permita, de uma coleção particular, inferir algo sobre outro caso particular.

A consequência lógica também pode ser entendida como um operador, um tipo especial de relação, qual seja, uma função, que associa conjuntos de fórmulas a conjuntos de fórmulas. Esta seria uma variação da Definição 1.3.1, conforme mostramos a seguir.

Definição 1.3.1' (Operador de consequência lógica Tarskiana)

Seja $\Im$ um conjunto não vazio, "$C(\Gamma)$" e "$C(\Delta)$", denotam, respectivamente, o conjunto de consequências de conjuntos Γ e Δ, para $\Gamma, \Delta \subseteq \Im$. Então um *operador de consequência Tarskiano sobre* $\Im$, ou uma *consequência lógica Tarskiana sobre* $\Im$, é uma função C: $\wp(\Im) \to \wp(\Im)$ tal que satisfaz as seguintes propriedades, para todo $\Gamma, \Delta \subseteq \Im$:

(T1') $\Gamma \subseteq C(\Gamma)$.

(T2') Se $\Gamma \subseteq \Delta$, então $C(\Gamma) \subseteq C(\Delta)$.

(T3') $C(C(\Gamma)) \subseteq C(\Gamma)$.

O operador de consequência lógica Tarskiano para um conjunto $\Im$ denotado por "$C_\Im$", pode ser associado ou identificado à relação de consequência lógica Tarskiana para $\Im$, denotado por "$R_\Im$", e vice-versa, de modo a garantir que um elemento x pertence ao conjunto de consequências de Γ, via $C_\Im$, se, e somente se, $\langle \Gamma, x \rangle \in R_\Im$. Tal associação pode ser estabelecida através das duas proposições abaixo, conforme citado por Gomes (2008, p 69). A expressão "sse" é uma abreviação para "se, e somente se,".

a: Relação de consequência associada ao operador de consequência

Dada uma função C: $\wp(\Im) \to \wp(\Im)$, a *relação de consequência associada a C*, denotada por "C_r", é a relação $R \subseteq \wp(\Im) \times \Im$ tal que, para todo $\Gamma \subseteq \Im$, tem-se que $\langle \Gamma, x \rangle \in R$ sse $x \in C(\Gamma)$.

b: Operador de consequência associado à relação de consequência

Dada uma relação $R \subseteq \wp(\Im) \times \Im$, o *operador de consequência associado a R*, denotado por "R_C" é a função C: $\wp(\Im) \to \wp(\Im)$ tal que, para todo $\Gamma \subseteq \Im$, tem-se que $R_C(\Gamma) = \{x \in \Gamma : \langle \Gamma, x \rangle \in R\}$.

Em áreas como a filosofia e a lógica, a serem seguidas neste trabalho, os elementos envolvidos na consequência lógica são geralmente considerados orações ou sentenças de uma dada linguagem.

Uma *sentença* pode ser definida como uma sequência finita apropriada de símbolos de uma linguagem. A definição daquilo que é apropriado é estabelecida em cada linguagem, através de sua gramática, especificamente, de suas regras de formação, como veremos no próximo capítulo. Tais regras devem distinguir, sintaticamente, as sentenças como elementos linguísticos que, semanticamente, são capazes de serem verdadeiros ou falsos.

Na lógica, uma linguagem é basicamente constituída de um alfabeto e de uma gramática. No alfabeto são estabelecidos os seus símbolos básicos. Na gramática são introduzidas as regras de formação, que discriminam certos agrupamentos de símbolos, dentre os quais encontram-se as sentenças.

Uma linguagem pode ser considerada parte de um sistema axiomático. De acordo com Shoenfield (1967, p. 02), um sistema axiomático é uma edificação constituída basicamente de conceitos básicos e derivados, de axiomas e teoremas. Eles podem ser associados a certos agrupamentos de símbolos ou ser construídos a partir de certos agrupamentos de símbolos estabelecidos na linguagem do sistema. Nós assumiremos a noção de sistema axiomático como base para a análise da consequência lógica.

Os conceitos básicos de um sistema axiomático devem ser tão claros e simples que podem ser entendidos sem uma definição precisa, diz Shoenfield (1967, p. 01). Os conceitos derivados são definidos a partir dos básicos. Pelo fato dos conceitos derivados serem associações dos básicos, acredita-se que a clareza inequívoca da base garanta a clareza inequívoca de todos os conceitos de um sistema.

Assim como os conceitos, as proposições de um sistema axiomático também são divididas em *básicas* e *derivadas*. As básicas são denominadas *axiomas* ou sentenças primitivas, conforme Tarski (1956c, p. 166). Em geral elas delimitam as propriedades elementares dos conceitos básicos. Nesse contexto, Shoenfield (1967, p. 01) as define como "... leis que nós percebemos como evidentes, considerando a natureza dos conceitos envolvidos". As teses derivadas são os *teoremas*, consequências do conjunto de axiomas, diz Tarski (1956c, p. 182). São derivados dos axiomas por meio de regras de inferência.

Tradicionalmente, os axiomas de um sistema axiomático são divididos em axiomas lógicos e axiomas não-lógicos. Na perspectiva dita clássica, inaugurada "oficialmente" por Aristóteles, os *axiomas lógicos* são aquelas teses básicas aceitas em qualquer sistema axiomático. Já os *axiomas não-lógicos*, ou postulados, são as teses básicas de um sistema específico.

Um axioma (ou um teorema) pode ser considerado sob duas perspectivas, diz Shoenfield (1967, p. 2-3):

> Pode ser visto como uma sentença, ou seja, como o objeto que aparece no papel quando escrevemos o axioma, ou como o significado de uma sentença, isto é, o fato expresso pelo axioma. [...] O estudo de axiomas e teoremas enquanto sentenças é chamado estudo *sintático* dos sistemas axiomáticos; o estudo do significado destas sentenças é chamado estudo *semântico* dos sistemas axiomáticos.

Conforme explica D'Ottaviano (1992, p. 66), a sintaxe é a dimensão combinatória de uma linguagem, encarada como puro jogo formal, sem significado. A dimensão semântica, por sua vez, leva em consideração objetos extra-linguísticos, aos quais as expressões da linguagem se referem, e o significado das mesmas.

Assim como as fórmulas, a consequência lógica também pode ser estudada via sintaxe ou via semântica. Em termos sintáticos, como veremos no próximo capítulo, a relação de consequência lógica entre um conjunto de premissas e uma conclusão é determinada através da manipulação de símbolos, a partir da aplicação adequada de regras lógicas adotadas. Já em termos semânticos, ela é determinada, dentre outras coisas, a partir da relação das expressões da linguagem com algum aspecto extralinguístico.

Em termos semânticos, o conceito usual básico na literatura utilizado para a determinação de uma consequência lógica é o de valor de verdade aplicado às premissas e à conclusão. A consequência lógica consiste na manutenção da verdade das premissas para a conclusão. A nossa proposta, a ser desenvolvida nos dois últimos capítulos, consiste na determinação de uma relação de consequência lógica a partir do valor informacional das sentenças.

Do ponto de vista semântico usual, costuma-se atribuir às premissas e conclusão algumas propriedades. Uma primeira é a de serem portadoras de verdade ou de falsidade. Elas podem ser verdadeiras ou falsas. Como a verdade e a falsidade são atributos de entidades linguísticas, uma segunda propriedade é a de pertencerem a uma linguagem. Em terceiro lugar, por serem portadores de verdade, elas devem ser indicativas (afirmar ou negar algo a respeito de algo). Isso exclui da análise da consequência lógica qualquer argumento constituído por entidades não-indicativas, como as interrogativas, imperativas e as interjeições. Em quarto lugar, em geral espera-se que possuam a propriedade de respeitar certos princípios, como o da impossibilidade de possuírem dois valores de verdade ao mesmo tempo, ou da exigência de possuírem pelo menos um valor de verdade.

Em suma, procuramos apresentar, neste capítulo, as três características centrais comumente atribuídas a uma definição adequada de consequência lógica. Expomos ainda uma definição clássica de consequência lógica, que assumiremos como referência para a classificação de diferentes tipos de consequência lógica. Explicitamos que assumiremos as fórmulas de uma linguagem como os elementos constituintes de uma relação de consequência lógica. Expomos a ideia de sistema axiomático, que possuem uma linguagem na qual é estabelecido, dentre outras coisas, o conjunto de sentenças. Tais sistemas são tomados como ponto de partida para as diferentes definições de consequência lógica expostas ao longo deste trabalho. No próximo capítulo tratamos do estudo sintático dos sistemas axiomáticos, a partir do qual expomos uma perspectiva sintática da consequência lógica.

Capítulo 2
Consequência lógica sintática

1 Apresentação

Neste capítulo tratamos do estudo sintático de um sistema axiomático, a fim de expor uma concepção de consequência lógica a partir de uma perspectiva sintática. Na próxima seção apresentamos uma definição de sistema formal lógico clássico, a partir do qual estabelecemos uma classificação entre os diferentes tipos de sistemas formais lógicos. Na terceira seção apresentamos a concepção de sistemas formais lógicos não-clássicos. Discutimos algumas características de certos sistemas em especial, visando mostrar as possíveis consequências para uma definição de consequência lógica a partir da perspectiva sintática, que chamaremos de *consequência lógica sintática*. Por fim, na última seção, expomos uma definição de consequência lógica sintática. Analisamos em que medida a perspectiva sintática satisfaz as características de necessidade, formalidade e anterioridade.

2 Sistemas formais lógicos

Um *sistema formal* pode ser caracterizado pelo fato de considerar unicamente a forma de seus objetos de discurso, sem atentar para outros fatores como o seu conteúdo ou sua referência. No caso destes objetos serem fórmulas, a natureza do sistema pode variar de acordo com as possibilidades de sua geração, podendo ser classificados, por exemplo, como dedutivos, indutivos ou abdutivos.

Neste trabalho abordaremos unicamente os sistemas formais dedutivos, cuja definição pode ser apresentada de uma maneira bastante geral. Feitosa & D'Ottaviano (2004, p. 71), por exemplo, definem um *sistema dedutivo* como um par formado por um conjunto qualquer não vazio L e um operador de consequência Tarskiano C sobre L. Em seguida, Feitosa & D'Ottaviano (2004, p. 72) apresentam uma definição de um tipo de sistemas dedutivos, construídos a partir de uma linguagem formal, denominados por eles de *sistemas lógicos*, nos quais L é o conjunto de fórmulas dessa linguagem. Devido aos nossos interesses neste trabalho, centraremos nossa análise unicamente nestes sistemas, que denominamos de *sistemas formais lógicos*, dos quais passamos a tratar a seguir.

Segundo Blanchette (2001, p. 117),

> Um sistema formal da lógica consiste de uma *linguagem formal* (uma coleção de fórmulas definidas somente em termos de sua sintaxe) rigorosamente especificada, juntamente com um *sistema dedutivo* (uma especificação daquelas séries de fórmulas que contam como *deduções* de fórmulas particulares a partir de conjunto de fórmulas).

Uma formalização de um sistema axiomático processa-se em quatro etapas, dizem Nagel e Newman (2003, p. 45-6):

> Primeiro, prepara-se um catálogo completo dos signos a serem usados no cálculo. Estes são o seu vocabulário [alfabeto]. Segundo, assentam-se as "Regras de Formação". Estas declaram quais das combinações dos signos do vocabulário são aceitáveis como "fórmulas" (de fato, como sentenças). Podemos considerar as regras como componentes da gramática do sistema. Terceiro, são estabelecidas as "Regras de Transformação". Elas descrevem a estrutura precisa das fórmulas a partir das quais são deriváveis outras fórmulas de dada estrutura. Estas regras são, com efeito, as regras de inferência. Finalmente, selecionam-se certas fórmulas como axiomas (ou como "fórmulas primitivas"). Elas servem de fundamento para o sistema inteiro.

Um bom sistema formal construído para um sistema axiomático deve apresentar algumas características, dentre elas: 1) representar adequadamente os conceitos do sistema através de símbolos de uma linguagem formal; 2) definir rigorosamente os agrupamentos significativos de símbolos, especialmente as suas fórmulas; 3) capturar os axiomas do sistema, representando-os por fórmulas da linguagem formal; 4) definir as regras que possibilitarão provar os teoremas do sistema; 5) aceitar como teoremas unicamente as sentenças que representam as orações do sistema axiomático que são consideradas as suas verdades.

Para Shoenfield (1967, p. 03) "... um sistema formal é a parte sintática de um sistema axiomático." Conforme lembram Carnielli e Epstein (2006, p. 286), Gödel determina que "Um sistema formal pode ser simplesmente definido como qualquer procedimento mecânico para produzir fórmulas." Para Tarski (1944, p. 346) um sistema formal, ou uma linguagem formalizada, é aquele cujas expressões de sua linguagem são exatamente especificadas referindo-se exclusivamente à sua forma.

Em sistemas totalmente formalizados, dizem Nagel e Newman (2003, p. 32), os axiomas e teoremas

> ... são "cadeias" (ou sequências finitamente longas) de marcas sem significado, construídas segundo regras para combinar os signos elementares do sistema em conjuntos maiores. Além disso, quando um sistema foi inteiramente formalizado, a derivação de teoremas a partir de postulados não é mais que a transformação (de acordo com regras) de um conjunto de tais "cadeias" em outro conjunto de "cadeias".

Abaixo apresentamos uma definição geral de um sistema formal lógico.

Definição 2.2.1 (Sistema formal lógico)

Um *sistema formal lógico* ou *teoria formal lógica*, denotado por "*F*", é composto pelas seguintes partes:

a: Uma linguagem formal.

b: Um conjunto de axiomas.

c: Um conjunto não vazio de regras de inferência.

Na linguagem são especificados o *alfabeto* (conjunto de símbolos básicos do sistema) e as *expressões bem formadas* (sequências finitas de símbolos do alfabeto). Uma das características das linguagens formalizadas, diz Tarski (1956c p. 166), "... é a delimitação precisa do "sentido" das suas expressões através de suas propriedades estruturais."

As *regras de inferência*, ou simplesmente regras, de um sistema formal apresentam a função de demarcar as deduções permitidas de fórmulas a partir de um dado conjunto de fórmulas. Cada regra de inferência, diz Shoenfield (1967, p. 04), "... determina que, sob certas condições, uma fórmula, chamada *conclusão* da regra, pode ser *inferida* a partir de outras fórmulas, denominadas *hipóteses* da regra." Uma regra é *finita* quando o seu conjunto de hipóteses é finito.

A escolha dos conceitos básicos, dos axiomas e das regras de inferência pode variar de acordo com os objetivos almejados na construção de um sistema (como o tipo de teoremas ou conceitos que se deseja derivar) ou depender inclusive das preferências estéticas e intuitivas de seu construtor.

A seguir definimos as noções de prova e teorema de um sistema formal lógico, seguindo definição usual na literatura.

Definição 2.2.2 (Prova em um sistema formal lógico)

a: Seja *F* um sistema formal lógico cujas regras de inferência são finitas. Uma *prova P em F* é uma sequência finita de fórmulas tal que cada uma delas ou é um axioma de *F* ou é a conclusão de uma regra de inferência de *F* cujas hipóteses precedem aquela fórmula na prova.

b: Uma fórmula φ *tem* uma prova em *F* sse houver uma prova em *F* na qual φ é a sua última fórmula.

Definição 2.2.3 (Teorema de um sistema formal lógico)

Uma fórmula φ é um *teorema* de *F* sse φ tem uma prova em *F*.

Uma vez estabelecida a noção geral de sistema formal lógico, a seguir definimos um tipo especial destes sistemas, que denominamos de sistemas formais lógicos clássicos. Eles servirão de base para classificarmos os sistemas em clássicos e não-clássicos. Seguimos a definição proposta por Shoenfield (1967).

Definição 2.2.4 (Alfabeto de uma linguagem de primeira ordem)

Uma *linguagem de primeira ordem*, denotada por "L^1", possui os seguintes símbolos, que constituem o seu *alfabeto*:

a: As variáveis "x", "y", "z", "w", "x'", "y'", "z'", "w'", ...

b: Para cada *n*, os símbolos de função *n*-ária e os símbolos de predicados *n*-ários.

c: Os símbolos: "¬", "∨", "∃", denominados, respectivamente de "negação", "disjunção" e "quantificador existencial".

d: Os símbolos "(" e ")", denominados de parênteses.

Para cada *n*, o número de símbolos de função *n*-ário pode ser zero ou não-zero, finito ou infinito. O mesmo pode ser dito para os símbolos de predicado, exceto que, entre os símbolos de predicado binários deve figurar o símbolo da igualdade "=".

Denotaremos a i-ésima função *n*-ária por "f^n_i" e o i-ésimo predicado *n*-ário por "P^n_i". As funções 0-árias são denominadas de *constantes individuais*. Os parênteses, símbolos de função e símbolos de predicado diferentes de "=" são denominados *símbolos não lógicos*. Os demais símbolos são denominados *símbolos lógicos*.

A seguir apresentamos algumas definições básicas de uma linguagem L^1, que constituem a sua gramática. Suprimiremos, quando possível a referência a L^1 nas frases. Assim, em vez de "expressão de L^1", escreveremos apenas "expressão".

Definição 2.2.5 (Regras de formação de expressões de uma linguagem)

a: Uma *expressão* é uma sequência finita de símbolos.

b: Um *termo* é um tipo de expressão caracterizado pela seguinte definição indutiva generalizada:

  i. Toda variável é um termo.

  ii. Se $t_1, ..., t_n$ são termos e f é uma função de grau n, então $ft_1 ... t_n$ é um termo.

  iii. Nada além de i e ii são termos.

c: Uma *fórmula atômica* é uma expressão do tipo $Pt_1 ... t_n$, em que $t_1, ..., t_n$ são termos e P é uma letra de predicado n-ária.

d: Uma *fórmula* é um tipo de expressão caracterizada pela seguinte definição indutiva generalizada:

  i: Toda fórmula atômica é fórmula.

  ii: Se **u** é uma fórmula, então $\neg$**u** é fórmula, denominada *negação* de **u** ou *fórmula negativa*.

  iii: Se **u** e **v** são fórmulas, então (**u** $\vee$ **v**) é fórmula, denominada *disjunção* de **u** e **v** ou *fórmula disjuntiva*.

  iv: Se **u** é fórmula e x uma variável, então $\exists x$**u** é fórmula, denominada *instanciação* de **u** ou *fórmula existencial.*

  v: Nada além de i a iv é fórmula.

A partir de agora, quando não gerar ambiguidade, eliminaremos os parêntesis das fórmulas. Utilizaremos as letras gregas "φ", "ψ" e "γ" como variáveis metalinguísticas que variam sobre fórmulas.

Definição 2.2.6 (Símbolos definidos)

a: "$\varphi \rightarrow \psi$" é uma abreviação de "$\neg\varphi \vee \psi$".

b: "$\varphi \wedge \psi$" é uma abreviação de "$\neg(\varphi \rightarrow \neg\psi)$".

c: "$\varphi \leftrightarrow \psi$" é uma abreviação de "$(\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)$".

d: “$\forall x\varphi$” é uma abreviação de “$\neg\exists x\neg\varphi$”.

Uma *fórmula definida*, em cada um dos casos acima, é uma expressão que resulta em uma fórmula quando os seus símbolos definidos são eliminados de acordo com as regras acima.

Dizemos que “$\varphi \rightarrow \psi$” é a *implicação material* de ψ por φ; “$\varphi \wedge \psi$” é a *conjunção* de φ e ψ; “$\varphi \leftrightarrow \psi$” é a *biimplicação* de φ e ψ e “$\forall x\varphi$” é a *generalização* de φ pelo x.

Apesar das fórmulas com os símbolos definidos não pertencerem diretamente à linguagem L^1 original, quando nos referirmos às fórmulas de L^1, também incluímos os quatro tipos considerados na definição acima.

Definição 2.2.7 (Definições adicionais para uma linguagem)

a: Uma *ocorrência* de uma variável x em uma fórmula φ é uma aparição de x em φ.

b: Uma ocorrência de uma variável x em uma fórmula φ é *ligada em* φ se x ocorre em uma parte de φ da forma $\exists x\psi$. Se não estiver ligada, dizemos que a ocorrência de x está *livre em* φ.

c: Uma *variável* x é *ligada* (*livre*) em φ se alguma ocorrência de x é ligada (*livre*) em φ.

d: Uma *sentença* é uma fórmula em que todas as suas variáveis são ligadas.

e: Uma *variável* x é *substituível por um termo* t *em uma fórmula* φ se, para cada variável y que ocorre em t, nenhuma parte de φ da forma $\exists y\ \psi$ contém uma ocorrência de x que é livre em φ.

Notação 2.2.8

a: A expressão “$\varphi_x[t]$” designa a expressão obtida a partir de φ pela substituição de cada ocorrência livre de x por t em φ, sendo que x é substituível por t em φ.

b: A expressão “$\varphi_x(t)$” designa a expressão obtida a partir de φ pela substituição de um número indefinido de ocorrências livres de x por t em φ, sendo que x é substituível por t em φ.

c: A expressão “$t_1 = t_2$” designa a expressão “$= t_1, t_2$”.

Definição 2.2.9 (Linguagem de primeira ordem)

Uma *linguagem de primeira ordem*, L^1, é uma linguagem na qual o seu alfabeto e fórmulas são como definidos acima.

Definida uma linguagem de um sistema formal lógico de primeira ordem clássico, a seguir introduzimos os seus axiomas.

Definição 2.2.10 (Esquemas de axiomas)

a: Um *axioma sentencial* é uma fórmula da forma $\neg\varphi \vee \varphi$.

b: Um *axioma da substituição* é uma fórmula da forma $\varphi_x[t] \rightarrow \exists x\varphi$.

c: Um *axioma da identidade* é uma fórmula da forma $x = x$.

d: Um *axioma da igualdade* é uma fórmula da forma $x_1 = y_1 \rightarrow \ldots \rightarrow x_n = y_n \rightarrow fx_1\ldots x_n = fy_1\ldots y_n$ ou da forma $x_1 = y_1 \rightarrow \ldots \rightarrow x_n = y_n \rightarrow (Px_1\ldots x_n \rightarrow Py_1\ldots y_n)$.

Definição 2.2.11 (Axioma lógico)

Um *axioma lógico* é uma fórmula que é um axioma sentencial, um axioma da substituição, um axioma da identidade ou um axioma da igualdade.

Por fim, introduzimos as regras de inferência de um sistema formal lógico de primeira ordem clássico, tal como introduzido neste trabalho.

Definição 2.2.12 (Regras de inferência)

a: A *regra da expansão* infere $\varphi \vee \psi$ a partir de ψ.

b: A *regra da contração* infere φ a partir de $\varphi \vee \varphi$.

c: A *regra da associação* infere $(\varphi \vee \psi) \vee \gamma$ a partir de $\varphi \vee (\psi \vee \gamma)$.

d: A *regra do corte* infere $\varphi \vee \psi$ a partir de $\gamma \vee \varphi$ e de $\neg\gamma \vee \psi$.

e: A *regra da introdução do quantificador existencial* infere $\exists x\varphi \rightarrow \psi$ a partir da fórmula $\varphi \rightarrow \psi$, se x não é livre em ψ.

Definição 2.2.13 (Sistema formal lógico de primeira ordem clássico)

Um *sistema formal lógico de primeira ordem clássico* ou *teoria de primeira ordem clássica*, denotada por “T^1”, é um sistema formal lógico tal que:

a: A linguagem de T^1, denotada por "$L(T^1)$", é uma linguagem de primeira ordem.

b: Os axiomas de T^1 são os axiomas lógicos de $L(T^1)$ e um conjunto de axiomas próprios, denominados *axiomas não-lógicos* de T^1.

c: As regras de T^1 são as regras da expansão, da contração, da associação, do corte e da introdução do quantificador existencial.

Definição 2.2.14 (Lógica de primeira ordem clássica)

Uma *lógica de primeira ordem clássica*, denotada por "LC^1", é uma teoria de primeira ordem clássica cujo conjunto de axiomas não-lógicos é vazio.

De acordo com as definições acima, toda lógica é uma teoria, que, por sua vez, pode ser pensada como uma lógica acrescida de um conjunto, possivelmente vazio, de axiomas não-lógicos. Os teoremas de uma teoria que não são teoremas da lógica subjacente são aqueles que dependem dos seus axiomas não-lógicos.

Devido aos nossos interesses neste trabalho, apresentamos a seguir uma abordagem particular de um sistema formal lógico, que denominamos de *Lógica Sentencial Clássica* ou, resumidamente, *LSC*. Construímos um sistema a partir do qual analisamos se uma fórmula de uma dada linguagem de um sistema formal lógico de primeira ordem clássico é um teorema do dado sistema utilizando unicamente as regras que dizem respeito aos conectivos lógicos, "$\neg$" e "$\vee$".

Definição 2.2.15 (Sistema formal lógico clássico sentencial)

Um *sistema formal lógico clássico sentencial* ou *teoria sentencial clássica*, denotada por "T^0" é assim constituída:

a: A linguagem de T^0, denotada por "$L(T^0)$", possui os símbolos "$\neg$", "$\vee$", "(", ")" e "A_1", "A_2", "A_3", ..., denominadas *letras sentenciais* ou *variáveis sentenciais*.

b: Uma *fórmula atômica* de $L(T^0)$ é uma letra sentencial isolada.

c: Uma *fórmula* é um tipo de expressão caracterizada pela seguinte definição indutiva generalizada:

i: Toda fórmula atômica é fórmula.

ii: Se **u** é uma fórmula, então $\neg$**u** é fórmula.

iii: Se **u** e **v** são fórmulas, então (**u** $\vee$ **v**) é fórmula.

iv: Nada além de i a iv é fórmula.

d: O conjunto de *axiomas* de T^0 é o subconjunto do conjunto de fórmulas de $L(T^0)$ constituído pelos axiomas sentenciais de $L(T^0)$, denominados *axiomas lógicos* de T^0, e por um conjunto de axiomas próprios de T^0, denominados *axiomas não-lógicos* de T^0.

e: As regras de T^0 são as regras da expansão, da contração, da associação e do corte.

Definição 2.2.16 (Lógica sentencial clássica)

A *lógica sentencial clássica*, denotada por "*LSC*", é uma teoria sentencial clássica cujo conjunto de axiomas não-lógicos é vazio.

A rigor, uma teoria sentencial clássica é um fragmento das teorias de primeira ordem clássicas. Nas teorias sentenciais clássicas são desconsiderados os quantificadores, símbolos de função e de predicados. Por isso, podemos considerar uma teoria sentencial clássica como um sistema formal clássico de ordem zero.

Pode-se mostrar que as fórmulas da linguagem de *LSC*, denotada por "*L*(*LSC*)", que possuem a seguinte forma são teoremas de *LSC*:

a: $\varphi \rightarrow \varphi$.

b: $(\varphi \wedge \psi) \rightarrow \varphi$.

c: $\varphi \rightarrow (\varphi \vee \psi)$.

d: $((\varphi \rightarrow \psi) \wedge \varphi) \rightarrow \psi$.

e: $((\varphi \rightarrow \psi) \wedge \neg\psi) \rightarrow \neg\varphi$.

f: $((\varphi \rightarrow \psi) \wedge (\psi \rightarrow \gamma)) \rightarrow (\varphi \rightarrow \gamma)$.

g: $\neg\neg\varphi \leftrightarrow \varphi$.

h: $(\varphi \wedge \psi) \rightarrow (\psi \wedge \varphi)$.

i: $(\varphi \rightarrow (\psi \rightarrow \gamma)) \rightarrow ((\varphi \rightarrow \psi) \rightarrow (\varphi \rightarrow \gamma))$.

j: $(\neg\varphi \rightarrow \neg\psi) \rightarrow ((\neg\varphi \rightarrow \psi) \rightarrow \varphi)$.

Existem sistemas formais lógicos em que os quantificadores podem ser aplicados também a predicados e funções, a predicados de predicados e a funções de funções e assim por diante. Eles são denominados de sistemas formais lógicos de ordem superior. Em uma teoria de ordem superior, seria possível dizer, por exemplo, que, "Para todo predicado P, P possui certa propriedade Q" ou que "Para alguma propriedade P, P se aplica a um dado indivíduo a".

Embora seja possível apresentar uma definição formal para os sistemas de ordem superior, não o faremos neste trabalho. Quando construídos conforme a definição de sistema formal clássico acima, um sistema formal lógico de n-ésima ordem, para $n > 1$ e natural, é considerado um *sistema formal lógico clássico*.

Dois sistemas formais podem apresentar notação, símbolos primitivos ou conjuntos de esquemas de axiomas ou de regras de inferência distintos e mesmo assim possuírem o mesmo poder de prova. A próxima definição trata destes sistemas.

Definição 2.2.17 (Sistemas formais lógicos dedutivamente equivalentes)
Sejam F^1 e F^2 dois sistemas formais lógicos e Teo(F^1) e Teo(F^2), respectivamente, os seus conjuntos de teoremas. Então F^1 e F^2 são *dedutivamente equivalentes* sse Teo(F^1) = Teo(F^2).

Dois sistemas formais são dedutivamente equivalentes ou *teoreticamente* equivalentes se provarem as mesmas fórmulas (ou possuem a mesma relação de consequência lógica sintática). Devem ser descontadas as diferenças de notação entre elas (os sinais gráficos que representam conceitos, como por exemplo, "$\neg$" e "$\sim$"), de símbolos primitivos (por exemplo, $\varphi \vee \psi =_{df} \neg\varphi \rightarrow \psi =_{df} (\varphi|\varphi) \mid (\psi|\psi)$ ou também que $\forall x\ \varphi =_{df} \neg\exists\neg x\ \varphi$), ou dos conjuntos de axiomas e regras de inferência.

Definição 2.2.18 (Sistema formal lógico clássico)
Um *sistema formal lógico clássico* é um sistema formal de n-ésima ordem clássico, T^n, para n natural, ou um sistema formal lógico dedutivamente equivalente a algum T^n.

Quando $n = 0$, temos uma teoria sentencial clássica. Quando $n = 1$, temos um sistema formal lógico de primeira ordem clássico. Quando $n > 1$ temos um sistema formal lógico de ordem superior clássico.

A essência dos sistemas formais ditos clássicos está na observação de três leis determinadas por Aristóteles e conhecidas como *Leis Básicas Aristotélicas do Pensamento*:

(L1) *Lei da (Não) Contradição*: algo não pode ser e não ser.

(L2) *Lei do Terceiro Excluído*: algo é ou não é.

(L3) *Lei da Identidade* ou *da Reflexividade da Identidade*: tudo é idêntico a si mesmo.

Essas três leis costumam, respectivamente, ser "representadas" nos sistemas formais lógicos pelas especificações das seguintes fórmulas:

(P1) $\neg(\varphi \wedge \neg\varphi)$.

(P2) $(\varphi \vee \neg\varphi)$.

(P3) $\forall x\,(x = x)$.

Na visão de Tarski (1944, p. 354), embora as fórmulas acima "representem" as leis do pensamento, elas não podem ser confundidas com as próprias leis, a fim de evitar contradições. As fórmulas são simplesmente expressões de uma linguagem, enquanto as leis são princípios que proíbem, por exemplo, que uma fórmula seja verdadeira e falsa ao mesmo tempo.

Para uma teoria ser clássica, todas as fórmulas da sua linguagem que forem especificações de (P1) – (P3) devem figurar entre os seus teoremas. Pode-se mostrar que todo sistema formal lógico clássico satisfaz este requisito.

No entanto, muitas teorias consideradas não-clássicas também possuem as especificações de (P1) – (P3) dentre seus teoremas. Na próxima seção tratamos destes sistemas, buscando explicitar as suas características próprias.

3 Sistemas formais lógicos não-clássicos

A história da lógica mostra que as leis aristotélicas básicas do pensamento foram, por mais de vinte séculos, consideradas tão intuitivas e óbvias que ninguém, em sã consciência, duvidaria de alguma delas. O sistema clássico aristotélico foi tido como o único apropriado para tratar de noções como a de consequência lógica. Qualquer sistema aceitável deveria seguir os cânones aristotélicos. Nessa visão absolutista, diz da Costa (1993, p. 39), "... há uma única lógica, que se pode formalizar de diversas maneiras distintas, mas todas se evidenciam equivalentes entre si, codificando as leis lógicas, únicas e imutáveis, que governam a realidade."

A lógica clássica, baseada nos cânones aristotélicos, no entanto, prova algumas fórmulas cuja validade intuitiva é duvidosa, tais como as que resultam nos chamados paradoxos da implicação material. Também permite algumas deduções intuitivamente condenáveis, tais como a de que de uma contradição tudo é consequência lógica ou que uma fórmula provada a partir de um conjunto vazio de hipóteses é consequência lógica de qualquer fórmula. À lógica clássica aristotélica, portanto, foram atribuídas algumas falhas na tentativa de capturar a noção de consequência lógica. Isso originou a sugestão de que essa lógica deveria ser modificada ou substituída.

Segundo Haack (2002, p. 208),

> As pressões para mudar os cálculos bivalentes clássicos, o sentencial e o de predicados, têm vindo de preocupações com a aparente inadequação do aparato clássico para representar os vários tipos de argumento informal, e sobre a interpretação e aplicação desse aparato.

Como mostra Haack (2002, p. 208-9) muitos pensadores reagiram a tais pressões propondo novas perspectivas. Dentre elas, foi sugerido reavaliar a forma ou a interpretação dos argumentos problemáticos ou dos símbolos da linguagem clássica, sem com isso modificar substancialmente o aparato clássico, essencialmente em sua parte sintática.

As reações dos defensores da lógica clássica não impediram o surgimento de novas lógicas. Seus conceitos primitivos e axiomas lógicos foram considerados intuitivamente claros e óbvios por muitos pensadores, que duvidavam da clareza ou obviedade de alguns princípios ou pressupostos clássicos.

Da Costa (1993, p. 16; 1997, p. 138), Mortari (2001, p. 356), Haack (2002, p. 29) classificam os sistemas não clássicos em dois tipos. Na definição abaixo procuramos capturar a distinção ontológica entre os sistemas não-clássicos.

Definição 2.3.1 (Sistemas formais lógicos complementares e heterodoxos)

Seja Teo(F) o conjunto de teoremas de um sistema formal lógico F qualquer; seja *Teo(F')* o conjunto de teoremas de um sistema formal lógico clássico *F'* e *Teo(F'')* conjunto de teoremas de um sistema formal lógico clássico *F''*. Então:

a. *F é um sistema formal lógico não-clássico complementar* sse Teo(F') $\subset$ Teo(F), para algum F', e Teo(F) $\neq$ Teo(F''), para todo F''.

b. *F é um sistema formal lógico não-clássico heterodoxo* sse Teo(F') $\not\subset$ Teo(F) , para algum F', e Teo(F) $\neq$ Teo(F''), para todo F''.

Uma teoria complementar, além de provar todos os teoremas de uma teoria clássica, também possui pelo menos um teorema próprio, não encontrado em qualquer teoria clássica. Isto requer que ela possua uma linguagem estendida da linguagem de uma teoria clássica, ou um conjunto maior de regras de inferência. Já uma teoria heterodoxa, cuja linguagem pode ser a mesma de um sistema formal lógico clássico, não possui, dentre os seus teoremas, alguns dos teoremas de alguma lógica clássica.

A divisão entre sistemas clássicos e não-clássicos não é absoluta, uma vez que nem sempre uma teoria pode ser classificada definitivamente. Como ilustra da Costa (1993, p. 16), a lógica intuicionista, por exemplo, pode ser encarada tanto como complementar quanto como heterodoxa, dependendo da perspectiva. Na classificação acima ala é considerada heterodoxa. Do mesmo modo, sistemas de lógica temporal podem ser construídos como um caso de uma lógica modal, portanto, complementar, ou como um sistema heterodoxo.

Segundo Haack (2002, p. 236), os lógicos complementares criticam a lógica clássica por ser muito restritiva. Ela prova teoremas e consequências corretas, mas não prova todas. Deixa de considerar, devido à pobreza de sua linguagem, argumentos intuitivamente válidos. Já os defensores das lógicas heterodoxas criticam a lógica clássica por ser demasiado permissiva. Os sistemas clássicos aceitam algumas consequências lógicas que não poderiam aceitar.

Dentre as lógicas complementares encontramos as lógicas epistêmicas, deônticas, imperativas e modais. As *lógicas epistêmicas* acrescentam à linguagem clássica os operadores “K”, a ser lido como “x sabe que”, e “B”, a ser lido como “x acredita que”. Já as *lógicas deônticas* acrescentam os operadores “O”, a ser lido como “deve ser o caso que” e “P”, a ser lido como “é permitido que”.

Com as *lógicas modais* busca-se representar argumentos que envolvem conceitos como os de necessidade e possibilidade (lógicas *aléticas*), obrigatório e permitido (lógicas *deônticas*), dentre outros.

No caso das lógicas modais aléticas, que discutimos a seguir, são acrescidos à linguagem clássica os operadores modais usualmente conhecidos como *necessário* (em geral representado pelo símbolo “□”) e *possível* (em geral representado pelo símbolo “◊”).

No começo do século XX, dois dos primeiros pensadores a tratar da lógica modal foram Peirce (1958) e Lewis (1918). Lewis (1918) foi um crítico da concepção de implicação implícita na lógica clássica, conhecida como implicação material. Nela, qualquer fórmula do tipo “se φ então ψ” pode ser reduzida à fórmula disjuntiva “não-φ ou ψ”. Isso origina um resultado paradoxal (no sentido de que vai além da opinião, do bom senso, que contraria a intuição). Tal resultado determina que a ocorrência de um dos dois disjuntivos de “não-φ ou ψ” garante que φ implica ψ. Assim, “A lua é de queijo”, por exemplo, implica “A bengala está no canto”. Na concepção de Lewis, uma fórmula φ implica uma fórmula ψ apenas quando for impossível que ocorra φ e não ocorra ψ.

Uma resposta a essa postura de Lewis (1918) poderia ser a de que há uma confusão entre implicação material e implicação lógica. Embora “A lua é de queijo” possa implicar materialmente “A bengala está no canto”, não a implica logicamente. Para tanto, deveria ocorrer a relação sugerida por Lewis. Essa resposta, no entanto, não serviria como defesa para outros paradoxos da implicação material, tais como os seguintes:

a: $(\varphi \rightarrow \psi) \vee (\psi \rightarrow \varphi)$.

b: $\varphi \rightarrow (\psi \rightarrow \varphi)$.

c: $\neg\varphi \rightarrow (\varphi \rightarrow \psi)$.

Qualquer instanciação das fórmulas acima podem ser provados em qualquer sistema formal lógico clássico. De acordo com a primeira delas, de duas fórmulas quaisquer, a primeira implica a segunda ou a segunda implica a primeira. No entanto, intuitivamente, observamos que “A lua é de queijo” não implica e não é implicada por “A bengala está no canto”. A segunda fórmula acima explicita que, se uma fórmula pode ser provada, então é implicada por qualquer fórmula. Contra a intuição, a segunda fórmula permitiria que “2 + 2 = 4”, por exemplo, fosse implicada por “2 + 2 = 5”. Já a terceira diz que, dada a negação de uma fórmula, a sua afirmação implica qualquer fórmula. De acordo com ela, “Érico Veríssimo não é escritor implica que se Erico Veríssimo é escritor, então a lua é de queijo”. Para evitar tais resultados, uma lógica deveria apresentar outra concepção de implicação.

Depois de alguns anos de estudo, Lewis e Langford (1932) desenvolveram alguns sistemas formais de lógica modal. No final dos anos 1950 começaram a ser apresentadas outras formalizações consideráveis para a lógica modal. Foram também associadas semânticas a sistemas modais, sugeridas por lógicos como Kripke (1963), capazes de fornecer um entendimento mais profundo de seus axiomas característicos.

Dentre as lógicas modais mais estudadas estão os sistemas K, T, S4, S5, que se diferenciam, dentre outros fatores, pelo conjunto de axiomas lógicos escolhido. A lógica sentencial modal S5, por exemplo, poderia ser definida como uma lógica sentencial acrescida de uma linguagem sentencial modal (uma linguagem sentencial acrescida de pelo menos um dos dois operadores modais aléticos acima citados), de fórmulas desta linguagem que são especificações dos esquemas de axiomas lógicos do esquema de sistema formal modal S5 e pelas regras de S5. Esse sistema contém todos os teoremas da lógica sentencial clássica e outros mais, tais como qualquer especificação da fórmula "$\Box\varphi \rightarrow \varphi$".

Ao contrário das lógicas complementares, que procuram ampliar a lógica clássica, as lógicas heterodoxas são rivais da lógica clássica. Dentre as lógicas heterodoxas encontram-se, por exemplo, certas lógicas temporais, intuicionistas, polivalentes e paraconsistentes.

Em determinados sistemas de *lógica temporal* não é possível provar sentenças do tipo $(\varphi \wedge \psi) \rightarrow (\psi \wedge \varphi)$. Considerando o fator tempo, há casos em que não vale a comutatividade da conjunção. A sentença "João entrou na sala e fechou a porta" pode não implicar a sentença "João fechou a porta e entrou na sala". Há lógicas temporais que podem ser consideras ampliativas, incluindo apenas um operador que representa instantes de tempo, a exemplo das lógicas modais.

Nas *lógicas intuicionistas* o problema não se encontra no tempo, mas em aspectos como a natureza de uma prova. O fundador do intuicionismo é o matemático holandês Brouwer, para o qual as provas que utilizam o princípio do terceiro excluído, como a maioria das provas por redução ao absurdo, por exemplo, não podem ser aceitas como genuínas ou legítimas na postura intuicionista. Provar que algo existe a partir de uma contradição resultante da suposição da sua inexistência não é uma prova aceitável. A prova da existência de uma entidade (como uma função ou uma relação) pressupõe a sua construção, ou a sua exibição.

Por considerar a intuição elemento fundamental na construção de provas, Brouwer era crítico em relação ao uso do formalismo lógico. Esse pensador não apresentou um sistema que formalizasse as ideias intuicionistas. Kolmogorov (1925) foi o primeiro a apresentar uma formalização de um sistema de caráter intuicionista. Mas foi o sistema de Heyting (1930), um discípulo de Brouwer, que ficou conhecido como a primeira formalização de uma lógica intuicionista.

Em sistemas intuicionistas não é possível provar, por exemplo, as fórmulas do tipo "$\varphi \vee \neg\varphi$" e "$\neg\neg\varphi \rightarrow \varphi$". Tampouco pode ser mostrado que φ é consequência lógica de $\neg\neg\varphi$.

Assim como a lógica intuicionista, as *lógicas polivalentes* também apresentam restrições ao princípio do terceiro excluído, mas por motivos distintos. Na primeira década do século XX, Peirce (1958) tratou deste tipo de sistemas. Mas um dos primeiros lógicos a propor uma lógica polivalente bem desenvolvida foi Łukasiewicz (1970), em um trabalho de 1920, aprimorada em trabalhos posteriores, alguns em parceira com outros lógicos, como Tarski e Lindenbaum. Uma apresentação dos principais resultados das álgebras dos cálculos polivalentes de Łukasiewicz pode ser encontrada em Cignoli, D'Ottaviano & Mundici (1992; 2000).

A crítica central de Łukasiewicz aos cânones clássicos refere-se ao princípio de que toda sentença deve ser ou verdadeira ou falsa. A existência de sentenças que não possuem valor de verdade ou que nem ela nem a sua negação podem ser provadas configuram exemplos de que esse princípio nem sempre está correto. A sentença "Haverá uma batalha naval amanhã", por sempre se referir ao futuro, não pode ser provada, nem ela, nem a sua negação.

Łukasiewicz (1970) propôs, inicialmente, um sistema a partir do pressuposto semântico de que uma sentença pode possuir um valor de verdade intermediário entre verdadeiro e falso. Ou seja, uma sentença pode apresentar três valores: verdadeira, falsa ou meio-verdadeira (ou então meio-falsa). Com base nesse sistema trivalente, esse pensador construiu cálculos n-valentes, com a possibilidade das sentenças possuírem n valores de verdade, para $3 \leq n \leq \aleph_0$.

Como enuncia D'Ottaviano (1992, p. 78), o cálculo sentencial bivalente de Łukasiewicz coincide com o cálculo sentencial clássico. Todas as lógicas n-valentes, com n finito ou infinito, são subsistemas do cálculo sentencial clássico, no sentido que os teoremas de uma lógica polivalente também são teoremas do cálculo sentencial clássico, mas nem sempre vale o inverso.

Por fim, as *lógicas paraconsistentes*, cuja criação se deve, em grande parte, ao brasileiro da Costa (1963; 1974), representam outra crítica à lógica clássica. Uma das principais características das lógicas em questão encontra-se na divergência com a concepção clássica de que uma teoria é inconsistente se, e somente se, for trivial. Uma *teoria é inconsistente* quando existe uma fórmula de sua linguagem tal que ela e a sua negação são seus teoremas. Uma *teoria é trivial* quando toda fórmula da teoria é teorema dela. As teorias paraconstistentes são aquelas que podem tratar da inconsistência, sem daí decorrer a trivialização da teoria.

Na Lógica Paraconsistente o problema de uma teoria não é a possibilidade de ela ser inconsistente, mas de ser trivial. Conforme da Costa (1958), do ponto de vista semântico e sintático, toda teoria é aceitável, desde que não seja trivial.

Nas lógicas paraconsistentes, o princípio da não-contradição não é necessariamente rejeitado. Mas, em toda lógica paraconsistente, de uma fórmula e sua negação não é possível, em geral, deduzir qualquer fórmula. Nessas lógicas, não é possível provar sentenças como "$\varphi \rightarrow (\neg \varphi \rightarrow \psi)$", válidas em todo sistema formal lógico clássico.

D'Ottaviano (1992, p. 84) explicita que da Costa propôs inicialmente um sistema de cálculos sentenciais C_n, para $1 \leq n \leq \omega$, que satisfazem as seguintes condições:

a: O princípio da não-contradição, na forma $\neg(\varphi \wedge \neg\varphi)$, não deveria ser válido em geral.

b: De duas premissas contraditórias φ e $\neg\varphi$, não deveríamos deduzir qualquer fórmula ψ.

c: Eles deveriam conter os mais importantes esquemas e regras da lógica clássica compatíveis com as duas primeiras condições.

Em suma, procuramos mostrar até o momento, neste capítulo, a existência de diferentes tipos de sistemas formais lógicos, com princípios e características radicalmente distintos.

Segundo Blanchette (2001, p. 117),

> Os primeiros sistemas formais, como o de Frege, por exemplo, foram construídos com a intenção de apresentar um modo para demonstrar rigorosamente relações de consequência lógica – isto é, demonstrar que afirmações particulares são de fato consequências lógicas de conjuntos particulares de afirmações. Vagamente falando, a intenção na construção desses sistemas era que o sistema pudesse incluir uma dedução de uma fórmula φ a partir de um conjunto Γ de fórmulas somente se φ fosse realmente uma consequência lógica de Γ. A dedutibilidade no sistema deveria ser um indicador confiável [necessário, mas não suficiente] da consequência lógica.

Na próxima seção analisamos as possíveis consequências desta multiplicidade de sistemas formais lógicos para a consequência lógica sintática.

4 Consequência lógica sintática

A consequência lógica sintática está embasada na ideia de obtenção de fórmulas a partir de conjuntos de fórmulas, uma pura manipulação de símbolos por meio de aplicações de regras de inferência, afirma Tarski (1956b, p. 63). É um cálculo realizado sem qualquer interpretação dos símbolos ou consideração a possíveis relações entre as expressões de uma linguagem e aos objetos ou situações a que tais expressões se referem.

Conforme Quine (1958, p. 09), uma das ideias centrais na perspectiva sintática da consequência lógica é a de que, se uma fórmula é consequência de outras, há uma prova mecânica daquela fórmula a partir destas, mesmo que ela ainda não tenha sido descoberta. A próxima definição explicita em que consiste uma dedução de uma fórmula ou de uma prova de uma fórmula a partir de um conjunto de fórmulas.

Definição 2.4.1 (Dedução de uma fórmula)

Sejam F um sistema formal lógico, cujas regras são finitas, ou seja, que possuem um conjunto finito de hipóteses, φ uma fórmula de $L(F)$ e Γ um conjunto de fórmulas de $L(F)$. Então uma *dedução* ou *prova de φ a partir de Γ* em F é uma sequência finita de fórmulas em que cada uma delas é um axioma de F, pertence a Γ ou é a conclusão de uma regra de inferência cujas hipóteses lhe precedem na prova e φ é a última fórmula da dedução.

Definição 2.4.2 (Consequência lógica sintática)

Uma fórmula φ é *consequência lógica sintática* de Γ em um sistema formal lógico F, que denotamos por "$\Gamma \vdash_F \varphi$", sse há uma dedução de φ a partir de Γ em F.

Quando Γ for o conjunto vazio, $\varnothing$, escreveremos $\vdash_F \varphi$ em vez de $\varnothing \vdash_F \varphi$. Se Γ for o conjunto $\{\varphi_1, ..., \varphi_n\}$, escreveremos $\varphi_1, ..., \varphi_n \vdash_F \varphi$ em vez de $\{\varphi_1, ..., \varphi_n\} \vdash_F \varphi$.

A partir da Definição 2.4.2 e da Definição 2.2.3 obtém-se que $\vdash_F \varphi \Leftrightarrow \varphi$ é teorema de F. As fórmulas que são consequência lógica do vazio em um sistema formal lógico são exatamente as mesmas que são teoremas desse sistema. Apesar disso, não podemos identificar as noções de consequência lógica sintática e de teorema. Enquanto um teorema é uma fórmula de uma teoria, a consequência lógica sintática é uma relação entre um dado conjunto de fórmulas e uma fórmula, que podem ou não ser teoremas.

A extensão da consequência sintática em uma teoria *T* é construída com base nos símbolos de sua linguagem, nos seus axiomas e regras de inferência, conforme ressalta Tarski (1956c, p. 181), por exemplo.

No plano conceitual ou intensional, uma proposta correta da consequência lógica deve apresentar as propriedades necessárias e suficientes atribuídas a ela intuitivamente. No plano extensional a adequação é definida pela equivalência das extensões da consequência nos âmbitos teórico e intuitivo. A seguir apresentamos algumas críticas à possibilidade da consequência lógica sintática ser adequada nos dois planos indicados.

A relação de consequência lógica sintática é usualmente definida a partir da concepção de prova, estabelecida como uma sequência *finita* de fórmulas. Isso pressupõe que as regras de inferência de qualquer sistema formal lógico que utilize esta definição devem possuir um conjunto finito de hipóteses. Esta característica gera um primeiro problema à adequação extensional dessa concepção, relacionada aos denominados ω-argumentos.

Tarski (1956d, p. 410) afirma ter construído uma teoria em que as sentenças do tipo "*n* possui a propriedade P", para *n* natural, são teoremas da teoria e a sentença "Todo número natural possui a propriedade P" não pode ser provada na teoria. É impossível provar que esta sentença geral seja consequência lógica daquelas sentenças particulares que, juntas com a sentença geral, constituem um ω-argumento, que consiste numa indução matemática. Para Tarski (1956d, p. 411),

> ... o conceito formalizado de consequência lógica, como geralmente usado pelos lógicos matemáticos [ou seja, o conceito sintático], não coincide com o conceito comum. Isso porque parece intuitivamente certo que a sentença 'Todo número natural possui a propriedade P' seque-se, no sentido usual, da totalidade de sentenças particulares '0 possui a propriedade P', '1 possui a propriedade P', '2 possui a propriedade P' e assim por diante.

Há tentativas de resolver este problema inserindo regras como a de indução infinita. Mas, diz Tarski (1956d, p. 411), embora essas regras permitam realizar a dedução desse tipo de argumentos, não se adequam à ideia de prova, dada a sua natureza infinitária. Na prática, jamais é possível obter a conclusão, dado que o número de hipóteses da regra é infinito. Uma proposta de consequência lógica baseada na ideia de provas infinitárias também apresentaria problemas.

Variações da regra, como as que agrupam todas as sentenças particulares em uma única sentença, poderiam torná-la finitária. Mas a sentença agrupadora, tal como "Todas as sentenças do tipo "*n* possui a propriedade P" são prováveis", não pertence à linguagem da teoria. Isso exigiria uma transição para a meta-teoria, o que implicaria em uma saída do sistema para a construção de uma prova.

Um segundo problema atribuído à consequência lógica sintática, especialmente à clássica, refere-se à monotonicidade. Há casos em que o acréscimo de hipóteses a um conjunto Γ de fórmulas pode desfazer uma consequência lógica entre uma fórmula φ e o dado conjunto Γ. A sentença "hoje tem aula de lógica" pode ser consequência lógica de "hoje é quinta feira" mas, se inserirmos a sentença "hoje é feriado", a consequência lógica pode ser desfeita.

A monotonicidade permite que se insiram fórmulas quaisquer em um conjunto de hipóteses, independente do conteúdo destas novas sentenças terem relação com o conteúdo da conclusão.

Há teorias não clássicas construídas para definir a consequência lógica a partir da relação de conteúdo relevante entre premissas e conclusão de um argumento. Dentre estes sistemas encontramos as lógicas da relevância, desenvolvidas inicialmente na década de 1920, por lógicos como Orlov (1928) e encontradas em Anderson & Belnap (1975).

Um terceiro problema da consequência lógica sintática refere-se à natureza intuitiva dos conceitos básicos de um sistema formal lógico. A explicitação da extensão da consequência lógica sintática depende e pode ser sugerida por meio da escolha de conceitos básicos, cuja compreensão é intuitivamente clara e transparente. A partir deles são definidos axiomas e regras apropriados para provar todos e somente os argumentos válidos ou os teoremas. Mas, dizem Nagel e Newman (2003, p. 28),

> ... a história do pensamento não tratou carinhosamente da doutrina do conhecimento intuitivo implícito na sugestão. Em certas áreas da pesquisa matemática onde suposições sobre coleções infinitas desempenham papéis centrais, aparecem contradições radicais, a despeito da clareza intuitiva das noções implicadas nas pressuposições e a despeito do caráter aparentemente consistente das construções intelectuais realizadas.

Há sistemas formais que, quando construídos, apresentaram axiomas aparentemente bem caracterizados, noções intuitivas claras e mesmo assim revelaram-se inconsistentes. Um exemplo é o sistema formal de Frege, assentado sobre noção supostamente clara e distinta de conjunto. Russell encontra um paradoxo nesse sistema, construindo o conhecido conjunto R de Russell, um conjunto que pertence a si mesmo se, e somente se, não pertence a si mesmo. Para Nagel e Newman (2003, p. 29)

> Esta contradição fatal resulta do uso não-crítico da noção aparentemente diáfana de classe. Outros paradoxos foram descobertos mais tarde, cada qual construído por meio de modos de raciocínio familiares e aparentemente cogentes [válidos]. Os matemáticos vieram a compreender que para o desenvolvimento de sistemas consistentes a familiaridade e a clareza intuitiva são caniços fracos para servir de apoio.

A clareza dos conceitos não é um elemento objetivo, podendo depender de fatores subjetivos. Com o desenvolvimento da lógica a partir do final do século XIX surgiram diversos sistemas capazes de provar um mesmo conjunto de fórmulas ou de argumentos válidos, como ocorreu com as lógicas sentenciais ou com as quantificacionais clássicas aristotélicas. Cada sistema apresenta um conjunto próprio de conceitos básicos (muitos deles considerados conceitos derivados em outros sistemas), axiomas e regras de inferência considerados óbvios e claros.

Embora esses sistemas clássicos divirjam quanto aos conceitos primitivos, todos apresentam a mesma extensão do conjunto dos teoremas ou da consequência lógica. O problema passa a ser realmente complicado quando começam a surgir sistemas diferentes igualmente apropriados e que pressupõem ou implicam uma modificação substancial da natureza do objeto investigado, como ocorre com os sistemas não-clássicos. Embora a noção de consequência lógica nestes sistemas, em geral, seja igualmente clássica, os pares de elementos pertencentes a cada relação de consequência lógica são diferentes.

A existência de diversas lógicas consideradas apropriadas, com extensões diferentes da consequência lógica, revela a falta de consenso sobre a sua extensão e, portanto, sobre a sua concepção. Em sua postura sintática, a lógica atualmente possui diversos paradigmas, alguns conflitantes, sobre a consequência lógica. O estabelecimento definitivo da concepção de consequência pressupõe um esclarecimento de quais destes paradigmas são aceitáveis e quais sistemas podem ser considerados uma lógica, conforme investigado por Haack (2002). Além disso, deve-se responder se há um único sistema considerado correto ou é possível a existência de vários sistemas capazes de capturar a noção de consequência lógica sintática.

Mesmo resolvendo o problema da natureza de uma lógica e da sua diversidade, haveria outro problema a ser resolvido sobre a concepção de consequência lógica sintática. Ao propor uma análise sintática da consequência lógica, espera-se que os sistemas formais sejam capazes de provar todos os argumentos ou teses reconhecidas intuitivamente como válidas, seja no âmbito estritamente lógico ou em alguma área específica do pensamento. As lógicas tratam apenas de teses ou de argumentos lógicos, provados apenas a partir dos axiomas lógicos. Não provam a validade de argumentos ou de teses referentes a domínios específicos, como a matemática, física ou ética.

O conteúdo de argumentos ou das fórmulas lógicas é absolutamente geral. Nada informam sobre aspectos particulares do universo. Os axiomas lógicos tratam apenas dos símbolos lógicos, que representam conceitos gerais, aplicáveis a qualquer domínio do conhecimento. É função dos axiomas não-lógicos das teorias estabelecer propriedades, apresentar informações sobre conceitos aplicáveis apenas a certos domínios particulares. Apenas as teorias com axiomas não-lógicos são capazes de provar argumentos ou fórmulas referentes a domínios particulares.

Para a concepção sintática de consequência lógica ser apropriada, é necessário que, a partir dela, sejam provadas todas e somente as fórmulas correspondentes às proposições consideradas intuitivamente válidas em um domínio. O intuito formalista sintático sempre foi o de demonstrar essa apropriação. No entanto, alguns resultados como os de Gödel provaram que essa apropriação é inalcançável para certos sistemas formais lógicos, considerados minimamente interessantes.

Em 1931 Gödel (1986) demonstrou, dentre outros resultados, que se T é uma teoria recursivamente axiomatizável consistente que "contém" a aritmética de primeira ordem, então existe alguma fórmula bem formada de sua linguagem a respeito dos números naturais tal que nem ela nem a sua negação podem ser provadas no sistema. No entanto, sabe-se que uma delas é verdadeira. Assim, independente dos conjuntos de axiomas ou de regras adotados, qualquer teoria do tipo em questão é incapaz de capturar todas as verdades da aritmética expressáveis na linguagem da aritmética de primeira ordem. Qualquer fragmento da aritmética capaz de tratar das operações básicas sobre os números naturais também apresenta essa característica.

Uma das propriedades fundamentais de uma teoria é a completude: o seu conjunto de teoremas deve ser igual ao seu conjunto de verdades. Ele deve provar todas e somente as verdades da teoria. Gödel mostrou, dizem Nagel e Newman (2003, p. 85-6),

> ... que há um número infinito de enunciados aritméticos verdadeiros que não se podem deduzir formalmente de qualquer conjunto dado de axiomas mediante um conjunto cerrado de regras de inferência. Segue-se que uma abordagem axiomática da teoria dos números, por exemplo, não pode esgotar o domínio da verdade aritmética. Segue-se, também, que o que entendemos por processo da prova matemática não coincide com a exploração de um método axiomático formalizado.

Para Gödel, lembram Carnielli e Epstein (2006, p. 286): "... formalistas consideraram a demonstrabilidade formal como sendo uma *análise* do conceito de verdade matemática e, por essa razão, não estavam naturalmente em posição de *distinguir* os dois."

As sentenças verdadeiras de um sistema formal lógico devem ser consequências dos seus axiomas. Se alguma delas não puder ser provada no sistema, a extensão da relação de consequência lógica também é incompleta.

As lógicas de primeira ordem não são afetadas pelos teoremas de Gödel e, assim, não sofrem as críticas direcionadas aos sistemas da aritmética. No entanto, a possibilidade de construção de diferentes lógicas igualmente apropriadas, que geram conjuntos de consequências lógicas distintas, como mostrado anteriormente, é um indicativo de que a consequência lógica sintática não satisfaz a característica da necessidade da consequência lógica. Isso por permitir que certas fórmulas sejam consequência lógica de um dado conjunto de fórmulas em uma lógica e que não o seja em outra lógica distinta. Assim, por exemplo, em certos sistemas formais lógicos, como os clássicos, qualquer fórmula de uma dada linguagem é consequência lógica sintática de uma fórmula do tipo "$\varphi \wedge \neg\varphi$", o que não acontece em sistemas como os paraconsistentes. Nos sistemas clássicos, em geral "$\varphi \vdash_T \psi \rightarrow \varphi$" e "$\neg\neg\varphi \vdash_T \varphi$", o que pode não acontecer em alguns sistemas não-clássicos. No entanto, para satisfazer a característica da necessidade, a relação de consequência lógica entre um dado conjunto de fórmulas e uma fórmula deveria se manter inalterada, independentemente do sistema formal lógico.

Essa mesma diferença entre sistemas também é um indicador de que a consequência lógica sintática tampouco satisfaz a característica da anterioridade. Para saber se uma fórmula é consequência lógica sintática de um conjunto de fórmulas em um sistema, é necessário antes saber quais são os seus axiomas, conceitos, regras de inferência.

Outro fator crítico da consequência lógica sintática é a possibilidade de existência de diferentes tipos de consequência lógica. A relação de consequência lógica subjacente em grande parte dos sistemas formais lógicos não clássicos é a Tarskiana. No entanto, existem definições de consequência lógicas não Tarskianas, que não satisfazem alguma das propriedades estabelecidas por Tarski. Esse é outro fator de que a consequência lógica sintática não satisfaz as características da necessidade e anterioridade.

Estes resultados negativos, entretanto, não tornam os sistemas formais lógicos e a definição de consequência lógica sintática absolutamente ineficazes ou inúteis. O conceito sintático de consequência lógica, diz Tarski (1956d, p. 413),

> ... provavelmente sempre terá uma significância decisiva para a construção prática de teorias dedutivas, como um instrumento que nos permite provar ou negar a possibilidade de provar sentenças particulares dessas teorias.

No entanto, conclui Tarski (1956d, p. 413), “O estudo formal sintático não revela as características essenciais da consequência lógica”. A sintaxe por si só parece ser insuficiente para a explicitação tanto da natureza da consequência lógica quanto de sua extensão. Para Etchemendy (1999, p. 157),

> Um sistema dedutivo fornece um modo de estudar a relação de consequência de uma linguagem, para provar resultados sobre ela, como a possibilidade de mecanizá-la. Mas não determina ou origina essa relação. Isso porque a questão referente ao fato de um sistema dedutivo particular para uma linguagem particular ser correto e completo é sempre uma questão sensata e, de fato, importante, de ser respondida.

A adequação da extensão da consequência sintática para uma linguagem sempre depende, em algum sentido, da sua contraparte semântica. Mostrar a completude de um sistema formal lógico, ou seja, que ele deduz tudo e somente tudo o que é válido, significa mostrar que a extensão da consequência lógica sintática é idêntica à extensão da consequência lógica semântica. Por trás desta afirmação está a ideia de que o estudo semântico parece capturar a concepção comum de consequência lógica. No próximo capítulo avaliamos se esta hipótese pode ser confirmada.

Em suma, neste capítulo procuramos expor e analisar as principais características e resultados da perspectiva sintática da consequência lógica. Neste estudo, a consequência lógica entre um dado conjunto de fórmulas e uma formula são produzidos a partir da mera manipulação de símbolos por meio da aplicação adequada de regras lógicas adotadas. Depois de apresentar a concepção de um sistema formal lógico clássico, expor e discutir a concepção de um sistema formal lógico não-clássico, apresentamos a concepção de consequência lógica sintática. O principal resultado deste capítulo foi o de que a perspectiva sintática não satisfaz as características da necessidade e anterioridade. Tal conclusão é baseada em dois motivos. O primeiro deles é a possibilidade de existência de diferentes tipos de sistemas formais lógicos com conjuntos de relações de consequência lógica distintos. O Segundo motivo encontra-se na possibilidade de existência de diferentes tipos de consequência lógica, clássicas e não-clássicas, consideradas igualmente apropriadas, de acordo com o objetivo almejado.

Capítulo 3
Consequência lógica semântica

1 Apresentação

Neste capítulo tratamos da consequência lógica a partir do estudo semântico de um sistema axiomático, que denominamos *consequência lógica semântica*. Na próxima seção apresentamos a abordagem modal da consequência lógica. Na terceira seção expomos a abordagem formal, dividida em duas versões, a substitucional e a interpretacional. Por fim, na última seção, apresentamos a abordagem modal/formal, uma espécie de síntese das duas abordagens anteriores. Para cada uma delas, procuramos expor as suas principais características e problemas, tomando como referência alguns exemplos de argumentos. Por fins didáticos, muitos destes argumentos são apresentados na Língua Portuguesa. Isso não gera prejuízo aos nossos propósitos neste capítulo, dado que a mesma análise poderia ser realizada com as traduções destes argumentos para uma linguagem formal, levando aos mesmos resultados. Em especial, visamos averiguar em que medida cada uma das perspectivas semânticas investigadas satisfaz as características da consequência lógica expostas no primeiro capítulo.

2 Abordagem modal

A *abordagem modal* se caracteriza por definir a relação de consequência lógica a partir do reconhecimento e análise das configurações possíveis do mundo ou das diferentes circunstâncias.

De acordo com Hanson (1997, p. 366), na perspectiva modal, uma sentença é consequência lógica de um dado conjunto de premissas quando for impossível que estas sejam verdadeiras e aquela seja falsa. Um argumento é legítimo quando qualquer circunstância concebível que tornasse verdadeiras as suas premissas tornaria também verdadeira a sua conclusão, diz Mates (1968, p. 04). Em outras palavras, quando for necessário que, dada a verdade das premissas, tem-se a verdade da conclusão.

Mesmo não sendo um adepto dessa abordagem, Etchemendy (1999, p. 81) concorda que “A característica mais importante da consequência lógica, como nós ordinariamente a entendemos, é uma relação modal mantida entre as premissas e a conclusão.” Segundo Hanson (1997, p. 366) a modalidade é frequentemente descrita em termos de verdade e falsidade em mundos possíveis (ou estados de coisas, ou circunstâncias).

Para Etchemendy (1999, p. 20), esta abordagem busca descrever, através de tabelas, utilizando, por exemplo, funções e tinta em papel, as diversas configurações do mundo, analisando a sua influência na determinação do valor de verdade de sentenças de uma linguagem. Estas representações, ou descrições linguísticas de mundos não-linguísticos, os mundos “efetivos”, são denominadas por Etchemendy (1999, p. 60) de *modelos*.

Na abordagem modal, uma sentença é *verdadeira em um dado modelo* quando for verdadeira no caso de o mundo ser como descrito por ele, diz Etchemendy (1999, p. 60). A verdade de uma sentença costuma ser definida com base na correspondência entre o que ela diz sobre o mundo e o que de fato ocorre nele. Com base na lógica clássica aristotélica, ilustramos essa concepção de verdade utilizando os seguintes argumentos (os termos de predicados são interpretados como de costume no Português, o nome “Érico Veríssimo” refere-se ao conhecido escritor gaúcho, “João” e “Maria” referem-se a indivíduos quaisquer):

Exemplo 3.2.1

a:	b:	c:
<u>Érico Veríssimo é gaúcho</u>	<u>João e Maria são casados</u>	<u>João e Maria são casados</u>
Érico Veríssimo é escritor	João é casado	João é casado com Maria

Para as duas sentenças do Exemplo 3.2.1a, existem quatro tipos de mundos possíveis, representados pelos seguintes modelos:

1. Érico Veríssimo é gaúcho e escritor.
2. Érico Veríssimo é gaúcho e não é escritor.
3. Érico Veríssimo não é gaúcho e é escritor.
4. Érico Veríssimo não é gaúcho e não é escritor.

Todo mundo possível, independente de sua constituição, em que Érico Veríssimo é gaúcho e escritor é do primeiro tipo, nesse caso, representado pelo primeiro modelo descrito acima. O mesmo pode ser dito sobre os outros tipos de mundos possíveis.

As sentenças do Exemplo 3.2.1a apresentam os seguintes valores de verdade nos modelos acima, pressupondo que o mundo seja como descrito em cada um deles:

	Sentenças	
Modelo	Érico Veríssimo é gaúcho	Érico Veríssimo é escritor
1	Verdadeira	Verdadeira
2	Verdadeira	Falsa
3	Falsa	Verdadeira
4	Falsa	Falsa

Se o mundo for como descrito pelo primeiro modelo, ambas as sentenças do Exemplo 3.2.1a são verdadeiras. Esse modelo representa o mundo *real*, aquele que de *fato* ocorre, desconsiderando o fator temporal (o instante em que Érico Veríssimo é gaúcho ou escritor). Por isso, dizemos que tais sentenças são *factualmente verdadeiras*, ou verdades de fato, e não apenas *hipoteticamente verdadeiras*, ou verdades hipotéticas.

No segundo modelo, ao contrário do primeiro, não há manutenção da verdade da premissa na conclusão. Se o mundo fosse como descrito pelo segundo modelo, a premissa seria verdadeira e a conclusão seria falsa. Logo, o argumento do Exemplo 3.2.1a é inválido, dada a possibilidade de a premissa ser verdadeira e a conclusão ser falsa.

O argumento do Exemplo 3.2.1b, por sua vez, é declarado válido na perspectiva modal. Se a sua premissa for verdadeira, tanto João quanto Maria são casados, o que garante a verdade da conclusão, independentemente de quem sejam João ou Maria. Se a conclusão fosse falsa, João não seria casado. A circunstância associada a este caso não seria logicamente possível, uma vez que, pela suposição da verdade da premissa, João é casado. Logo, João seria e não seria casado, contrariando o princípio da não-contradição.

O argumento do Exemplo 3.2.1c é inválido na visão modal. A sua premissa informa apenas que João é casado, mas não delimita o cônjuge. É possível a existência de um mundo no qual João seja casado com alguém distinto de Maria. De acordo com esse mundo e, portanto, no modelo que o representa, a premissa seria verdadeira e a conclusão falsa.

A definição das circunstâncias concebíveis ou dos mundos logicamente possíveis pode variar. Nos três argumentos acima utilizamos como critério os princípios da lógica clássica aristotélica. Para o argumento do Exemplo 3.2.1b é impossível conceber uma circunstância que torne verdadeira a premissa e falsa a conclusão, dado o seu caráter contraditório. Esse talvez não fosse o caso se adotássemos como base uma lógica não-clássica.

A especificação dos cânones lógicos adequados representa um primeiro problema para a perspectiva modal. A quantidade de mundos possíveis em uma lógica que rejeita algum dos princípios da lógica clássica aristotélica pode variar. Uma lógica que permita contradições, por exemplo, também permitiria modelos que representam mundos em que algo ocorre e não ocorre ao mesmo tempo. Isso poderia resultar em definição distinta da noção de validade de argumentos e efetuar mudanças na extensão da relação de consequência lógica. Tal variação da definição de validade pode suscitar dúvidas sobre a capacidade desta perspectiva de capturar a noção comum de consequência lógica, dada a suposição intuitiva da existência de apenas uma extensão desta noção.

Além da definição dos princípios lógicos, diz Sher (1996, p. 660), a abordagem modal deve explicitar todas as inter-relações possíveis e necessárias entre indivíduos, propriedades e relações denotadas pelos termos de uma dada linguagem, a fim de estabelecer os mundos possíveis a serem considerados. Ela poderia exigir, por exemplo, que a interseção da extensão dos termos branco e vermelho seja vazia ou que se um objeto é grande não pode ser pequeno sob o mesmo aspecto. Em linguagens razoavelmente ricas, tal conjunto de informações pode ser enorme e dificilmente poderia ser organizado.

Sher (1996, p. 558) denomina a perspectiva modal de *semântica metafísica*. Isso porque, para ela, tal abordagem deve estabelecer como os objetos referidos pelos termos de uma linguagem podem se apresentar em um mundo, o que, consequentemente, determina os tipos de modelos e a verdade das sentenças de um argumento. Esse estabelecimento depende da natureza, da essência desses objetos. Considerando o argumento do Exemplo 3.2.1a, deve-se determinar a possibilidade da existência de um mundo em que Érico Veríssimo seja um mineral, uma cidade ou um maratonista profissional. Sher (1996) procura mostrar que tais respostas, quando apresentadas, nem sempre o são de modo satisfatório. Se a definição dos mundos possíveis é obscura, conclui-se que a definição de consequência lógica também o deve ser, uma vez que depende dela.

A perspectiva modal mantém fixo o significado dos termos das premissas e conclusão. A linguagem deve ser completamente interpretada, diz Sher (1996, p. 659). Por isso, ainda com base nos princípios da lógica clássica aristotélica, são considerados válidos, erroneamente, os seguintes argumentos, cujos termos possuem o significado usual em Português:

Exemplo 3.2.2

a:	b:	c:
Érico Veríssimo é gaúcho	$(\exists x)$ x é água	Érico Veríssimo é gaúcho
Érico Veríssimo é brasileiro	$(\exists x)$ x é H_2O	Pelé é escritor ou não o é.

A suposição da invalidade do argumento do Exemplo 3.2.2a pressuporia a possibilidade de um mundo em que a premissa fosse verdadeira e a conclusão falsa. Nesse mundo, Érico Veríssimo teria nascido no Brasil (pois no significado de "gaúcho" encontramos a característica de ser brasileiro) e não teria nascido no Brasil. Esse argumento permite existirem apenas três tipos de mundos possíveis:

1. Érico Veríssimo é gaúcho e brasileiro.
2. Érico Veríssimo não é gaúcho e é brasileiro.
3. Érico Veríssimo não é gaúcho e não é brasileiro.

A quarta possibilidade é excluída por levar a uma contradição. Como em todos os modelos logicamente possíveis há a manutenção da verdade, a conclusão do Exemplo 3.2.2a é uma consequência lógica da sua premissa, segundo a abordagem modal.

Ao declarar o Exemplo 3.2.2a como válido por este método, a abordagem modal desconsidera a característica da formalidade. A sua validade é garantida graças ao conteúdo das suas sentenças, uma condição extra-lógica. Argumentos com a mesma estrutura podem apresentar a premissa verdadeira e a conclusão falsa, sendo intuitivamente considerados inválidos, como é o caso do seguinte argumento, uma variação do Exemplo 3.2.2a:

Exemplo 3.2.3

Érico Veríssimo é gaúcho.

Érico Veríssimo é paulista.

O argumento acima pode ser utilizado para ilustrar outro problema desta abordagem. Segundo Etchemendy (1999, p. 60), ela não determina como o valor de verdade de sentenças seria afetado se o objeto de discurso fosse inexistente. Caso Érico Veríssimo não tivesse existido, a sentença "Érico Veríssimo é gaúcho" teria seu valor de verdade indeterminado. Isso impediria discernir a respeito da manutenção de verdade em certas situações, o que impossibilitaria a determinação da consequência lógica entre sentenças desse tipo.

Outro problema da perspectiva em análise, causado pela definição do significado ou da referência dos termos, é apresentado por Kripke (1980) com base no Exemplo 3.2.2b. Se considerarmos que a água é essencialmente H_2O, não é possível este argumento apresentar conclusão falsa e premissa verdadeira. Em um mundo assim, algo seria e não seria água ao mesmo tempo. A validade do argumento do Exemplo 3.2.2b depende, além do conteúdo das sentenças, de um conhecimento prévio, qual seja, a identificação do significado ou da referência de "água" e "H_2O". Ou seja, a abordagem modal desconsidera a anterioridade da consequência lógica.

A validade do Exemplo 3.2.2c depende do significado dos termos "ou" e "não", mostrando outra vez que a abordagem modal não satisfaz a característica da anterioridade. Por causa do significado desses termos, e devido aos cânones da lógica clássica aristotélica, não há mundo possível que falsifique a conclusão. Logo, é impossível que a premissa seja verdadeira e a conclusão falsa.

O Exemplo 3.2.2c revela outro problema, não apenas da perspectiva modal, mas também de outras perspectivas. A validade desse argumento independe do conteúdo da premissa, totalmente irrelevante para a conclusão. Há lógicas, como as da relevância, que procuram estabelecer um sistema explicativo da consequência lógica considerando a relevância das hipóteses para a conclusão. Do ponto de vista estritamente modal, a relevância deve indicar outros pressupostos para a definição de mundos possíveis, a fim de impedir a validade desse argumento. Isso também modificaria, no mínimo, o conjunto da relação de consequência lógica.

De acordo com Etchemendy (1999, p. 25), o valor desta abordagem

> ... não está na análise das noções de verdade e consequência lógica, ou na análise da verdade analítica ou necessária. Em vez disso, esta abordagem nos proporciona uma perspicaz estrutura para caracterizar as regras semânticas que governam nosso uso da linguagem sob investigação. Deveria ser vista como um

> método de abordagem do estudo empírico da linguagem e não uma tentativa de analisar qualquer dos conceitos empregados nesta tarefa.

Em suma, o elemento modal, isoladamente, parece não ser adequado para capturar a noção de consequência lógica, tanto do ponto de vista da sua extensão quanto da satisfação das suas características, especialmente as de formalidade e anterioridade. Assim, não pode ser adequada para definir essa noção.

A seguir apresentamos uma segunda abordagem semântica da consequência lógica, buscando analisar em que medida ela satisfaz os requisitos básicos para uma boa definição da consequência lógica.

3 Abordagem formal

A *abordagem formal* não pode ser confundida com o *âmbito* formal, como tratado no capítulo anterior sobre sistemas formais lógicos. Ambos se identificam por considerar a forma das expressões como elemento central na análise de noções como a de consequência lógica, mas se distinguem quanto à sua natureza. Uma abordagem formal pode ser apresentada tanto no âmbito formal quanto no informal. Nesta seção analisamos informalmente duas versões da abordagem formal da noção de consequência lógica: substitucional e interpretacional.

A abordagem formal assemelha-se à abordagem modal por adotar a manutenção da verdade e a análise de uma totalidade de circunstâncias, geralmente definidas a partir de um conjunto distinto de símbolos lógicos, como elementos definidores da consequência lógica. Mas se distingue da abordagem modal pelo tipo de análise realizada. Enquanto a modal se preocupa em investigar mundos logicamente possíveis, a formal busca considerar especificações da estrutura das sentenças, considerando as diferentes significações dos termos não-lógicos de uma linguagem. Na modal, modificam-se os mundos e mantêm-se inalterados os elementos das sentenças, permanecendo fixo o significado dos termos envolvidos. Na formal, são modificados os elementos das sentenças (de modo a não alterar a sua estrutura) e mantém-se inalterado o mundo, variando o conteúdo dos termos não-lógicos.

Apresentações informais da abordagem formal, encontradas em manuais de lógica, diz Hanson (1997, p. 366-7), determinam que:

> A conclusão de um argumento é consequência lógica de suas premissas caso nenhum argumento com a mesma forma lógica tenha premissas verdadeiras e conclusão falsa. Certamente isso pressupõe algum modo de especificar a forma lógica, que é geralmente feito com base na distinção entre termos lógicos e não-lógicos de uma linguagem.

De acordo com Mates (1968, p. 17),

> O conceito de forma lógica é de capital importância no estudo que fazemos. Por isso mesmo, requer análise cuidadosa. É necessário dispormos de critérios práticos para decidir qual a forma lógica de uma dada sentença e quais as formas que só tenham sentenças analíticas como especificações. Infelizmente, a irregularidade das linguagens naturais torna a obtenção desses critérios difícil, se não impossível; só se pode esperar obtê-los nas linguagens artificiais.

O fato da identificação da forma lógica das expressões de uma linguagem formal poder ser bem estabelecida não significa que ela possa ser encontrada facilmente. Mates (1968, p. 19) admite que a determinação da forma "... varia de acordo com os vocábulos lógicos; infelizmente, a questão de saber quais vocábulos devem ser considerados lógicos, e quais não, envolve certo grau de arbítrio."

Na perspectiva formal costuma-se definir que uma sentença de uma linguagem é verdadeira em um dado modelo ("interpretação" dos termos não-lógicos da linguagem) se o que ela disse sobre o mundo na interpretação sugerida é, de fato, o caso, diz Etchemendy (1999, p. 61). Assim, em um mundo em que Érico Veríssimo tivesse nascido na Argentina, e que "Érico Veríssimo" e "brasileiro" tivessem o significado habitual, a sentença "Érico Veríssimo é brasileiro" seria falsa.

Com base nestas considerações apresentamos a seguir, em linhas gerais, duas versões da abordagem formal, acima enunciadas. Comecemos com a análise da versão substitucional da consequência lógica.

3.1 Versão substitucional

Na *versão substitucional*, que pode ser encontrada em Bolzano (1837) e Quine (1958), a noção de consequência lógica é definida a partir de substituições apropriadas de termos nas premissas e conclusão. Para uma conclusão ser consequência lógica de um dado conjunto de premissas, cada substituição permitida dos seus termos mantém aquela verdadeira, a partir da verdade destas.

Segundo Hanson (1997, p. 367), a versão substitucional

> ... determina que a relação de consequência lógica se mantém sempre que nenhuma *substituição* de termos não-lógicos por termos não-lógicos (respeitadas as categorias gramaticais) produz premissas verdadeiras e conclusão falsa. (De acordo com essa versão, argumentos com a mesma forma são aqueles que podem ser obtidos um a partir do outro por certo tipo de substituição.)

Na perspectiva substitucional, um método para a análise da validade de uma sentença φ é assim descrito por Etchemendy (1999, p. 32):

> Primeiro, introduzimos um estoque de variáveis para cada categoria gramatical. Em seguida substituímos cada *termo variável* [símbolo não-lógico] na sentença φ com uma *variável* de tipo apropriado, assegurando que múltiplas ocorrências de um termo recebam a mesma variável e termos distintos recebam variáveis distintas. O resultado desta operação é uma função sentencial *φ'* contendo somente expressões que ocorrem em Π, o conjunto de termos fixos escolhido. Consideremos, então, a coleção de instâncias de substituição de *φ'* – a coleção de sentenças que resultam de *φ'* pela inserção de expressões categorialmente apropriadas no lugar das variáveis. Se cada elemento deste conjunto for verdadeiro então φ é julgada logicamente verdadeira com respeito à atual escolha de termos fixos; se um ou mais for falso, então φ não é logicamente verdadeira com respeito àquela escolha.

De acordo com esse método, para uma fórmula φ ser consequência lógica de um conjunto Γ de fórmulas, com respeito a uma escolha particular de símbolos lógicos, toda instância de substituição φ' e Γ' de φ e de Γ, respectivamente, deve preservar a verdade de φ' a partir de Γ'.

Para ilustrar o método substitucional, consideremos o argumento do Exemplo 3.3.1a abaixo e sua instância exposta no Exemplo 3.3.1a', escritos em uma contração da língua portuguesa, que denominamos de linguagem *C*. As únicas constantes de nomes de *C* são "Érico Veríssimo" e "Pelé" e as únicas constantes de predicado são "Gaúcho" e "Brasileiro":

Exemplo 3.3.1

a:	b:
<u>Érico Veríssimo é brasileiro.</u>	<u>Pelé é brasileiro.</u>
Érico Veríssimo é gaúcho.	Pelé é gaúcho.

Todas as sentenças dos dois argumentos acima possuem a mesma forma. Ela consiste de um sujeito (nome próprio), o verbo ser na terceira pessoa do singular (considerada a única constante lógica de *C*) e um predicado (adjetivo de naturalidade).

Se as constantes não-lógicas tiverem seu significado ou a sua extensão usual, ambas as sentenças do Exemplo 3.3.1a são verdadeiras, pressupondo que Érico Veríssimo é brasileiro e gaúcho. Já o argumento do Exemplo 3.3.1a' não preserva a verdade, uma vez que a sua premissa é verdadeira e a sua conclusão é falsa nesta interpretação de *C*. Como o Exemplo 3.3.1a' é uma instância do Exemplo 3.3.1a, o argumento é, apropriadamente, considerado inválido pelo método substitucional.

O Exemplo 3.3.1a' ilustra que a manutenção da verdade no argumento do Exemplo 3.3.1a não se deve apenas à sua forma, mas também ao fato de Érico Veríssimo satisfazer a ambos os predicados desse argumento. A verdade das sentenças do Exemplo 3.3.1a é resultado da adequação do seu conteúdo aos fatos, um fator extra-lógico.

A versão substitucional apresenta alguns pontos falhos, dentre eles os que denominamos *problema dos termos lógicos* e *problema da complexidade da linguagem*.

Expomos o problema dos termos lógicos partindo dos seguintes argumentos, que parecem ser intuitivamente válidos:

Exemplo 3.3.2

a:	b:
Todo homem é mortal.	João sabe que Pedro é inocente.
Sócrates é homem.	João acredita que Pedro é inocente.
Sócrates é mortal.	

Estes dois argumentos são variações, respectivamente, de exemplos expostos por Quine (1958, p. 07) e Etchemendy (1999, p. 29). A sua validade na perspectiva substitucional depende dos termos "Todo", "é", "sabe" e "acredita" serem escolhidos como lógicos. Isso garantiria a verdade da conclusão, dada a verdade das premissas, em qualquer troca permitida dos demais símbolos, considerados não-lógicos. Caso os termos "Todo" e "sabe" não fossem considerados lógicos, teríamos as seguintes variantes, cujas premissas poderiam ser verdadeiras e a conclusão ser falsa:

Exemplo 3.3.3

a:	b:
Algum homem é mortal.	João deseja que Pedro seja inocente.
Sócrates é homem.	João acredita que Pedro é inocente.
Sócrates é mortal.	

A definição da validade de um argumento em uma linguagem na versão substitucional é baseada na escolha dos seus termos lógicos. Dependendo desta escolha, tal versão pode ser muito permissiva, considerando indevidamente válidos alguns argumentos, ou muito restritiva, considerando inválidos argumentos que não deveria. Essa dependência da especificação da consequência lógica a partir dos símbolos lógicos, somada à dificuldade de defini-los, resulta no que denominamos *problema dos termos lógicos*.

Conforme ressalta Etchemendy (1999, p. 30), em uma linguagem minimamente complexa, com um conjunto vazio de termos lógicos, nada é consequência de nada. Sempre seria possível substituir os termos não-lógicos de um argumento de tal modo a tornar verdadeiras as suas premissas e falsa a conclusão. Por outro lado, se todo termo fosse considerado lógico, a consequência lógica dependeria do valor de verdade factual das sentenças da linguagem. Nesse caso, a noção de consequência lógica seria reduzida à consequência material. Se o sinal de implicação for um símbolo não-lógico, ilustra Sher (1996, p. 663), a regra de *Modus Ponens*, intuitivamente válida, seria inválida na versão substitucional.

Há pensadores que acreditam ser possível construir um conjunto apropriado de termos lógicos capazes de capturar a noção de consequência lógica. Quine (1970), por exemplo, sugere que eles sejam os símbolos lógicos da lógica quantificacional clássica, como explicitados na Definição 2.2.13. Porém, tal escolha não é totalmente fundamentada, dada a carência de uma definição clara e precisa da noção de termo lógico.

A versão substitucional apresenta também o *problema da complexidade da linguagem*. A definição de consequência lógica, ou pelo menos da sua extensão, depende do tipo e da quantidade de termos não-lógicos de uma linguagem. Esta versão viola o que Etchemendy (1999, p. 31) denomina de *condição da persistência*: a propriedade de *não ser consequência lógica* deve persistir em *contrações* da linguagem e a propriedade de *ser consequência lógica* deve persistir nas *expansões*. A violação da persistência, ou problema da complexidade da linguagem, permite que argumentos inválidos passem a ser válidos em contrações da linguagem (eliminação de termos não-lógicos) e argumentos válidos passem a ser inválidos quando ela for estendida (inserindo novos termos não-lógicos).

A seguir apresentamos o problema da complexidade da linguagem relativo à versão substitucional, visando mostrar que ele impede tal versão de satisfazer as características da formalidade e da necessidade da consequência lógica. Para alcançarmos este objetivo, partimos da seguinte proposição, sugerida por Tarski (1956d, p. 415):

> (*F*): *Se, nas sentenças da classe* Γ *e na sentença* φ*, as constantes - exceto as constantes puramente lógicas – são substituídas por quaisquer outras constantes (signos iguais sendo substituídos, em toda ocorrência, por signos iguais), e se nós denotamos a classe de sentenças assim obtidas de* Γ *por* 'Γ'' *e a sentença obtida de* φ *por* 'φ'' *então a sentença* φ' *deve ser verdadeira dado apenas que todas as sentenças de* Γ' *sejam verdadeiras.*

Na proposta Tarskiana, (F) deve ser uma condição necessária e suficiente para uma definição adequada da consequência lógica. Assim, um argumento é válido se, e somente se, satisfizer (F).

Na versão substitucional, (F) é uma condição necessária para uma sentença φ ser uma consequência lógica da classe Γ. Se ela não for satisfeita para um φ e um Γ então a versão substitucional, apropriadamente, não considera φ uma consequência lógica de Γ. Dizer que (F) não é satisfeita para um argumento *A* significa dizer que há algum argumento *A'* com a mesma forma de *A* que não preserva a verdade. Nestas condições, pelas características da formalidade e da necessidade, *A* realmente não poderia ser considerado válido.

Se um argumento *A* preservar a verdade e houver um argumento com a mesma forma que não a preserva, conclui-se que a manutenção da verdade em *A* não depende apenas da forma das suas premissas e conclusão. Isso porque cada Γ' e φ' possui exatamente a mesma forma de Γ e φ. Como a forma é a mesma em toda variação de Γ e φ, se a preservação da verdade não for mantida em alguma variação, essa mudança é devida a algum outro fator além da forma. Para satisfazer a característica da formalidade, o argumento deve ser considerado inválido, como determina a condição (F). É isso o que ocorre com o Exemplo 3.3.1a acima e sua variante, exposta no Exemplo 3.3.1a'.

O Exemplo 3.3.1a também ilustra a parte da contração do problema da complexidade da linguagem para a versão substitucional. Se a linguagem *C*, subjacente ao argumento do Exemplo 3.3.1a fosse contraída, eliminando dela o nome "Pelé", aquele argumento passaria a ser válido nessa nova linguagem. Intuitivamente, porém, o Exemplo 3.3.1a não pode ser considerado válido, independentemente da linguagem adotada. Este seria um caso, diz Tarski (1956d, p. 415) em que

> ... a sentença φ não se segue no sentido ordinário das sentenças da classe Γ embora a condição (F) seja satisfeita. Esta condição pode realmente ser satisfeita somente porque a linguagem com a qual estamos tratando não possui uma quantidade suficiente de constantes extra-lógicas.

Para ilustrar a parte da expansão do problema da complexidade da linguagem, utilizaremos os dois argumentos abaixo, expressos em *C*, a linguagem subjacente ao Exemplo 3.3.1.

Exemplo 3.3.4

a:	a’
<u>Érico Veríssimo é gaúcho.</u>	<u>Érico Veríssimo é gaúcho.</u>
Érico Veríssimo é brasileiro.	Érico Veríssimo é paulista.

O argumento do Exemplo 3.3.4a é válido segundo *C* porque preserva a verdade em qualquer substituição possível em *C*. Uma expansão de *C* na qual fosse inserido o termo “paulista”, tornaria esse argumento inválido. Nessa linguagem expandida, o Exemplo 3.3.4a’ seria uma instanciação do Exemplo 3.3.4a. Como o argumento do Exemplo 3.3.4a’ não preserva a verdade, o Exemplo 3.3.4a não seria válido em *C* expandida. Isso mostra que a condição (F), embora necessária, não é suficiente para definir a consequência lógica, conforme enuncia Tarski (1956d, p. 415).

O Exemplo 3.3.4a mostra que a versão substitucional não satisfaz a característica da formalidade. A manutenção da verdade deste argumento não se deve à forma das sentenças, mas a um elemento extra-lógico, a relação geográfica entre São Paulo e Brasil.

Tal versão tampouco satisfaz a característica da necessidade porque a definição da noção de consequência lógica depende de outro fator extra-lógico, a complexidade da linguagem assumida. O número de circunstâncias ou mundos possíveis analisados depende do número de termos não-lógicos da linguagem. Nos argumentos acima, por exemplo, desconsideram-se, de acordo com *C*, a maioria dos mundos cujos indivíduos não sejam apenas Érico Veríssimo e Pelé.

A principal desvantagem da postura substitucional, diz Mates (1968, p. 18),

> ... reside no pressuposto de que a linguagem contém um nome ou descrição para cada objeto e um predicado correspondente a cada propriedade (Com efeito, objetos e propriedades sem nome poderiam ser precisamente os que viessem a tornar falsa a matriz.) Por vários motivos, aquele pressuposto não é admissível.

Se a intuição servir de guia, argumentos como o do Exemplo 3.3.4a não podem ser considerados válidos em uma definição apropriada de consequência lógica. Isso nos leva a concluir que (F), quando satisfeita, nem sempre distingue corretamente os argumentos válidos na versão substitucional. Nessa versão, (F) não é suficiente para a sentença "φ é consequência lógica de Γ". Somente o seria no caso em que todo objeto possível pudesse ter uma designação na linguagem. Entretanto, diz Tarski (1956d, p. 416),

> ... esta hipótese é fictícia e jamais pode ser satisfeita. Nós devemos procurar algum meio de representar as intenções da condição (F) que seja inteiramente independente desta hipótese fictícia.

A seguir apresentamos esse meio sugerido por pensadores como o próprio Tarski. Pretendemos avaliar a capacidade da proposta Tarskiana na solução dos problemas expostos acima e de satisfazer as características da consequência lógica.

3.2 Versão interpretacional

A *versão interpretacional* assemelha-se à versão substitucional por priorizar a característica da formalidade, distinguir os termos lógicos de uma linguagem e estabelecer a manutenção da verdade como elemento definidor da consequência lógica. Distingue-se dela por interpretar (atribuir "significados") em vez de substituir os símbolos não-lógicos nas sentenças, mantendo inalterado o seu conjunto de símbolos.

Segundo Hanson (1997, p. 367)

> ... [a versão interpretacional] determina que a relação de consequência lógica é mantida sempre que nenhuma *interpretação* dos termos não-lógicos em um argumento produz premissas verdadeiras e conclusão falsa. De acordo com essa versão, um argumento com a mesma forma que um dado argumento é o próprio argumento, porém com todos, uma parte ou nenhum, de seus termos não-lógicos interpretados diferentemente. Dizer que um termo não-lógico é interpretado diferentemente significa simplesmente que lhe é atribuída uma intensão diferente (ou pelo menos uma extensão diferente).

Atribuir uma intensão ou uma extensão diferente a um termo significa alterar, respectivamente, uma ou mais propriedades que o caracterizam ou um ou mais indivíduos que o satisfazem.

Um dos principais expoentes da versão interpretacional é Tarski (1956). Seu objetivo é propor uma noção de consequência lógica que satisfaça as condições de formalidade e necessidade expressas na condição (F) acima enunciada.

A proposta Tarskiana, em algum sentido, é também substitucional. Porém, em vez de substituir os termos não-lógicos de um argumento por outros termos da linguagem, Tarski os substitui por variáveis não pertencentes à linguagem. Tal substituição implica em interpretações ou re-interpretações dos termos não-lógicos. A definição de consequência lógica Tarskiana depende dos objetos do universo e não da complexidade de uma linguagem.

Tarski (1956c) apresenta uma concepção de verdade, aplicada às linguagens de primeira ordem, equivalentes às da Definição 2.2.9. Elas são assumidas como linguagens objeto, para cujas fórmulas é definido um valor de verdade, a partir do qual é definida a noção de consequência lógica. A proposta Tarskiana é baseada em algumas definições formais prévias, como as seguintes, que reproduzimos de uma maneira livre.

Definição 3.3.5

a: Uma *função sentencial de uma linguagem L* é uma fórmula de L com um conjunto de variáveis livres.

b: Uma *sentença de L* é uma função sentencial de L sem variáveis livres.

c: Uma *sequência S de objetos satisfaz uma função sentencial α*, denotado por "Sat(S, α)" sse os objetos de S cumprem os requisitos da função sentencial.

d: Uma *sentença φ de L é verdadeira segundo uma linguagem ("interpretação")* sse toda sequência S satisfaz a função sentencial correspondente a φ em uma metalinguagem correspondente.

e: Uma *sequência S é um modelo de um conjunto Γ* de fórmulas de L sse S satisfaz todas as fórmulas de Γ. Em particular, quando $\Gamma = \{\varphi\}$, dizemos que S é um modelo da fórmula φ.

Em resumo, a satisfação é uma relação mantida entre uma sequência de objetos e uma função sentencial. Intuitivamente, é a relação que se apresenta, por exemplo, entre a pessoa Érico Veríssimo e a função sentencial “x é gaúcho”, mas que não se mantém entre ela e o indivíduo Pelé. No primeiro caso a relação se mantém porque Érico Veríssimo é gaúcho. Já no segundo não há a relação porque Pelé não é gaúcho. A relação de satisfação entre uma sequência de objetos e uma função sentencial depende do significado dos símbolos não-lógicos de cada linguagem *L*.

Pode-se mostrar, a partir das definições acima, conforme o faz o próprio Tarski (1956c, p. 198), os seguintes resultados:

a: Se uma sequência S satisfaz uma sentença φ, então, para toda S’, S’ satisfaz φ.

b: Se uma sequência S não satisfaz uma sentença φ, então, para toda S’, S’ não satisfaz φ.

c: φ não é verdadeira (é falsa) sse, para toda sequência S, S não satisfaz φ.

A partir destas definições, Tarski (1956d, p. 417) define a concepção de consequência lógica, conforme o fazemos a seguir.

Definição 3.3.6 (Consequência lógica de Tarski)

Uma sentença φ é *consequência lógica de Tarski* de um conjunto Γ de sentenças, denotado por “CLT(Γ, φ)” sse, para toda sequência S, se S é modelo de Γ, então S é modelo de φ.

Utilizaremos a expressão “~~CLT~~(Γ, φ)” para denotar que “φ não é consequência lógica de Tarski de Γ”.

Enquanto o valor de verdade de uma sentença é definido a partir de uma linguagem subjacente, a definição de CLT independe de qualquer linguagem subjacente. A validade de um argumento é estabelecida independentemente de qualquer interpretação dos símbolos não-lógicos.

Como é de se esperar, pode-se facilmente mostrar que CLT é um operador de consequência Tarskiano, conforme a Definição 1.3.1.

A abordagem interpretacional, tal como exposta por Tarski, não apresenta o problema da persistência, discutido na seção anterior: se CLT(Γ, φ) em *L*, então CLT(Γ, φ)

em qualquer expansão de *L*; se ~~CLT~~(Γ, φ) em *L*, então ~~CLT~~(Γ, φ) em qualquer contração de *L*. Isso pode ser observado a partir da seguinte argumentação.

Para Tarski (1956d, p. 417), CLT é completamente independente do significado dos símbolos não-lógicos das premissas e conclusão. Além disso, a validade de um argumento em uma linguagem *L* independe da quantidade de nomes de *L*. Ao substituir os nomes de indivíduos ou de predicados (ou qualquer outro símbolo não-lógico) por variáveis de uma metalinguagem correspondente, onde serão interpretados estes símbolos, CLT é definida a partir da análise de todo objeto ou de toda relação entre objetos, através do seu significado na metalinguagem correspondente, e não de seu nome na linguagem objeto. Assim, no caso da expansão de uma linguagem, é indiferente se o nome de um objeto ou de um predicado figura nela. Quando lhe é inserido um nome de indivíduo, por exemplo, é como se ele já tivesse sido analisado, porque o objeto que ele nomeia já foi considerado. Nesse sentido, diz Etchemendy (1999, p. 37) "... a satisfação nos dispõe de todos os nomes *possíveis* que poderiam ser incorporados na linguagem."

O que foi dito acima também vale para o caso da contração de uma linguagem. A eliminação de um nome de indivíduo na linguagem não o exime de ser considerado na metalinguagem. Assim, se a relação de consequência lógica não é satisfeita por um conjunto de fórmulas e uma fórmula em uma dada linguagem, tampouco se manterá em uma contração dela.

De acordo com Tarski (1956d, p. 417),

> Parece-me que qualquer um que entende o conteúdo da definição acima deve admitir que ela concorda muito bem com o uso comum. Isso se torna ainda mais claro a partir de seus vários resultados. Em particular, pode ser provado, com base nesta definição, que toda consequência de sentenças verdadeiras deve ser verdadeira.

Em outras palavras, é possível mostrar que: se CLT (Γ,φ), então se as sentenças de Γ são verdadeiras, φ é verdadeira. A prova desta proposição pode ser apresentada a partir das definições sugeridas por Tarski (1956c; 1956d). Etchemendy (1999, p. 86), ao reproduzir uma versão desta prova, afirma que, aparentemente, ela independe da escolha dos símbolos lógicos ou de mundos possíveis. Ao não fazer referência a nenhum conjunto de símbolos lógicos ou mundos possíveis na demonstração da manutenção da verdade da CLT, pode-se inferir, equivocadamente, que CLT garante a manutenção da verdade em qualquer conjunto de símbolos lógicos ou para qualquer mundo possível.

No entanto, diz Etchemendy (1999, p. 86), a demonstração da manutenção da verdade mostra apenas que se CLT(Γ, φ) e se todas as sentenças de Γ são verdadeiras então φ também é verdadeira. Isso equivale a dizer que as seguintes condições são incompatíveis:

(P) CLT(Γ, φ) (para algum conjunto Π fixo de símbolos lógicos).

(Q) Todos os elementos de Γ são verdadeiros.

(R) φ é falsa.

Isso, porém, não implica mostrar que:

Se

(P) CLT(Γ, φ) (para algum conjunto Π fixo de símbolos lógicos),

então as duas condições seguintes são incompatíveis:

(Q) Todos os elementos de Γ são verdadeiros.

(R) φ é falsa.

Os dois conjuntos de sentenças acima constituem a Falácia de Tarski (FT):

(FT)

(F1) Necessariamente (Se P e Q então não-R).

(F2) Se P então necessariamente (Se Q então não-R).

Para Etchemendy (1999, p. 85), inferir (F2) de (F1) significa cometer a *Falácia de Tarski*. A diferença entre (F1) e (F2) está na consideração do conjunto de símbolos lógicos ou na análise de um mundo particular, em geral o mundo *real*. Enquanto (F1) trata, em cada caso particular, de uma circunstância específica, (F2) generaliza uma escolha particular.

Não há como provar qualquer garantia da manutenção da verdade com base em CLT, diz Etchemendy (1999, p. 92). Independentemente de quais símbolos lógicos sejam escolhidos ou de qual mundo possível tenha sido analisado, não há como assegurar que, assumindo que CLT(Γ, φ) em um caso particular, o será em todos.

Para mostrar a invalidade de (FT), basta exibir um “mundo” em que (F1) seja verdadeira e (F2) seja falsa. Como visto acima, (F1) é verdadeira. O argumento do

Exemplo 3.3.7a, construído em uma linguagem *L*, alterada de modo a considerar lógicos todos os seus símbolos, e sua interpretação, exposta no Exemplo 3.3.7a', falsifica (F2):

Exemplo 3.3.7

a:	a'
Ga	Érico Veríssimo é gaúcho.
Ea	Érico Veríssimo é escritor.

Como ambas as sentenças do Exemplo 3.3.7a são verdadeiras de fato, CLT(Ga, Ea) na linguagem subjacente. Mas, intuitivamente, esta não é consequência genuína, pois a premissa não sustenta a conclusão. Em uma escolha diferente do conjunto de símbolos lógicos, o Exemplo 3.3.7a pode possuir a premissa verdadeira e a conclusão falsa. Isso ocorre quando o termo "Érico Veríssimo" é não-lógico e interpretado como Pelé. O argumento do Exemplo 3.3.7a é um caso em que CLT(Γ, φ) (segundo o conjunto de símbolos lógicos escolhidos) embora não seja considerada intuitivamente uma consequência lógica. Há possibilidade de a premissa ser verdadeira e da conclusão ser falsa. Isso ilustra uma circunstância em que (F2) é falsa e (F1) é verdadeira.

Para Etchemendy (1999, p. 93), a abordagem de Tarski iguala a consequência lógica à preservação da verdade de um conjunto de argumentos, delimitado pela escolha dos símbolos lógicos. Não há garantias de que a verdade das premissas assegura a verdade da conclusão. Mas, segue Etchemendy (1999, p. 93),

> Um argumento logicamente válido deve, pelo menos, ser capaz de justificar a sua conclusão. Deve ser possível saber que a conclusão é verdadeira com base no conhecimento de que o argumento é válido e que as premissas são verdadeiras. Esta é uma característica dos argumentos logicamente válidos que, mesmo aqueles mais céticos com relação às noções modais, reconhecem como essencial. ... [Com a abordagem de Tarski] Não temos segurança que argumentos que satisfazem a definição [de consequência lógica] serão capazes de justificar sua conclusão e, assim, nenhuma garantia de que eles serão genuinamente válidos. A Falácia de Tarski obscurece esta omissão, ao estabelecer que argumentos declarados válidos de fato garantem a preservação da verdade.

A crítica acima se baseia essencialmente na generalização errônea encontrada em (F2), provocada por (FT). De uma escolha particular de um conjunto de símbolos lógicos, conclui-se a manutenção da verdade para qualquer escolha de símbolos lógicos.

(FT) realiza também uma generalização sobre o conjunto de mundos possíveis. A manutenção da verdade toma como base um mundo possível, geralmente o *real*, a partir do qual a Falácia infere implicitamente a manutenção da verdade para qualquer mundo possível, sem prévio exame.

A fim de evitar generalizações indevidas, talvez a prova da manutenção da verdade deva ser restringida: CLT(Γ, φ) segundo um conjunto Π de símbolos lógicos implica a manutenção da verdade apenas segundo Π. Assim, a manutenção da verdade dependeria apenas da forma do argumento, especificamente dos símbolos lógicos presentes nele.

Etchemendy (1999, p. 101-6) procura provar que a suposição acima tampouco pode ser justificada. Mesmo nesse caso restrito, a concepção Tarskiana pode considerar válidos argumentos devido à dependência de aspectos extra-lógicos. O argumento do Exemplo 3.3.8a abaixo, construído em uma linguagem *L* de um sistema formal lógico de primeira ordem clássico, acrescida de "P" e "Q" no conjunto de símbolos lógicos, e sua "interpretação", exposta no Exemplo 3.3.8a' ilustram esse problema.

Exemplo 3.3.8

a:	a':
Pa____	Dilma é presidente do Brasil.
Qa	Dilma tem mais de 35 anos.

Como todos os presidentes do Brasil até o momento foram maiores de 35 anos, para toda sequência S, se S for modelo da premissa, então S é modelo da conclusão. A validade do Exemplo 3.3.8a depende essencialmente de um fato histórico, político, considerado um fator extra-lógico. No mundo logicamente possível cuja presidência do Brasil seja ocupada por alguém com menos de 35 anos, a conclusão pode ser falsa e a premissa verdadeira.

O Exemplo 3.3.8a ilustra que CLT é inadequada extensionalmente mesmo em situações em que é mantido o conjunto de símbolos lógicos na demonstração da manutenção da verdade. Isso por desconsiderar a característica da necessidade, que exige a análise de todo mundo possível e não a penas o *real*, critica Etchemendy (1999, p. 109).

Ao permitir que um argumento seja declarado válido com base em fatores extra-lógicos, CLT também desconsidera a anterioridade. Para estabelecer a validade do Exemplo

3.3.8a é preciso conhecer a história política do Brasil. A validade pressupõe uma verificação anterior do mundo *real*. Com isto, diz Etchemendy (1999, p. 108), a definição de CLT, seja em sua versão restrita (mantendo um mesmo conjunto de símbolos lógicos) ou na geral (generalizando a prova para todo conjunto de símbolos lógicos), falha em capturar a noção intuitiva de consequência lógica, uma vez que ele deve ser totalmente independente do que de fato ocorre no mundo.

O próprio Tarski (1956d, p. 418-9) admite que a definição de CLT para uma linguagem depende da escolha de seus símbolos lógicos. No caso extremo, se todos os símbolos de uma linguagem forem lógicos, o conceito de CLT formal coincidiria com o conceito material. Assim, a fórmula φ seria consequência lógica Tarskiana de Γ se pelo menos uma sentença de Γ fosse falsa ou φ fosse verdadeira.

Na posição de Tarski (1956d, p. 418),

> ... a escolha dos símbolos lógicos não é totalmente arbitrária. Se, por exemplo, considerarmos não-lógicos os símbolos de implicação ou do quantificador universal, então nossa definição do conceito de consequência poderia conduzir a resultados que obviamente contradiriam o uso ordinário. Por outro lado, desconheço qualquer critério objetivo que nos permita dividir precisamente os símbolos lógicos e não-lógicos.

Talvez o problema das generalizações indevidas na demonstração da manutenção da verdade e da influência de fatores extra-lógicos pudesse ser resolvido se houvesse um conjunto Π de símbolos genuinamente lógicos. Nesse conjunto não poderiam figurar os nomes de indivíduos ou de predicados, para evitar que a validade seja influenciada pela natureza dos indivíduos ou das propriedades envolvidas nos argumentos. Isso invalidaria argumentos como os do Exemplo 3.3.7a e do Exemplo 3.3.8a.

Muitos pensadores assumem que esse conjunto seja constituído pelos símbolos lógicos explicitados anteriormente na Definição 2.2.4, quais sejam: "x_i", "y_i", "z_i", "w_i", "$\neg$", "$\vee$", "$\exists$" e "$=$". Este conjunto pode variar, dependendo da escolha dos conectivos ou quantificadores assumidos como básicos. Neste caso, o esquema de argumento do Exemplo 3.3.9 abaixo, escrito em uma linguagem desta lógica, ilustra que a proposta Tarskiana ainda sofreria influências de fatores extra-lógicos como o tamanho do universo:

Exemplo 3.3.9

$\underline{\exists x_u \ (x_u = x_u \lor x_u \neq x_u) \qquad\qquad}$

$\exists x_1 \ldots \exists x_n \prod^{n}_{i,,j=1} (x_i \neq x_j)$, para $i < j$.

Uma n-instanciação do Exemplo 3.3.9 é um argumento cuja conclusão possui n repetições de quantificador existencial. Então:

Grau da instanciação	*Conclusão*
2	$\exists x_1 \exists x_2 \ (x_1 \neq x_2)$
3	$\exists x_1 \ \exists x_2 \ \exists x_3 \ (x_1 \neq x_2 \land x_2 \neq x_3 \land x_1 \neq x_3)$
4	$\exists x_1 \ \exists x_2 \ \exists x_3 \ \exists x_4 \ (x_1 \neq x_2 \land x_2 \neq x_3 \land x_3 = x_4 \land x_1 \neq x_3 \land x_1 \neq x_4 \land x_2 = x_4)$

A premissa de uma n-instanciação do Exemplo 3.3.9 determina a existência de pelo menos um objeto igual ou diferente de si mesmo, enquanto a conclusão afirma existirem pelo menos n objetos no universo.

Intuitivamente, qualquer n-instanciação do Exemplo 3.3.9 parece ser inválida. Não há garantias da existência de pelo menos n objetos a partir do fato de existirem objetos iguais ou diferentes de si mesmos. Assim, para a abordagem de Tarski ser adequada, deve considerar inválida qualquer n-instanciação do Exemplo 3.3.9. Deve haver uma sequência que satisfaça a função sentencial correspondente à premissa na metalinguagem e não satisfaça a função sentencial correspondente à conclusão. Como todos os símbolos nas sentenças do Exemplo 3.3.9 são considerados genuinamente lógicos, essa sequência deve satisfazer a própria premissa e não satisfazer a conclusão do Exemplo 3.3.9. Com isso, a sua conclusão deve ser falsa e a sua premissa verdadeira, considerando que uma sequência satisfaz uma fórmula se, e somente se, toda sequência a satisfaz.

Para uma sequência S de objetos satisfazer a premissa do Exemplo 3.3.9 deve haver pelo menos uma sequência S' que difere de S no máximo em seu u-ésimo elemento tal que o u-ésimo elemento de S' satisfaz um dos dois disjuntivos que compõem a sentença em questão, o que realmente deve ocorrer para qualquer elemento. Nos termos da proposta de Tarski, para S não satisfazer a conclusão do Exemplo 3.3.9, toda sequência S' deve possuir objetos i e j tais que $1 \leq i,j \leq n$, em que $i < j$, o i-ésimo e o j-ésimo objetos de S' tais que x_i

$= x_j$. Para isso ocorrer, é necessário que o universo tenha menos do que *n* objetos. Se possuir acima de *n* elementos é possível encontrar uma sequência na qual nenhum dos *n* primeiros elementos sejam iguais entre si, o que a permitirá satisfazer a conclusão, tornando-a verdadeira.

Pode-se notar que se o universo for finito, contendo *k* objetos (para *k* natural), a proposta Tarskiana considera válida toda *k*-instanciação do Exemplo 3.3.9, para $k > n$, e inválidas as demais. Apenas a infinitude do universo validaria toda *n*-instanciação. Assim, a adequação de CLT depende do tamanho do universo, um fator extra-lógico, conclui Etchemendy (1999, p. 111).

Para Etchemendy (1999), um dos principais problemas da proposta Tarskiana está na redução da relação de consequência lógica entre sentenças à relação de consequência lógica das suas funções sentenciais correspondentes em uma metalinguagem. Isso permite que a consequência lógica seja determinada por fatores extra-lógicos, referentes ao mundo em que a relação de satisfação das funções sentenciais está sendo analisada. Hanson (1997) também credita a esta característica o principal problema da inadequação da proposta de Tarski e de outras versões da abordagem formal. Segundo Hanson (1997, p 370),

> Esta inadequação é claramente devida ao fato de que ambas a versões da abordagem formal fazem uso da verdade *simpliciter*, isto, é a verdade no mundo real, em vez de verdade em todo mundo possível ou verdade em todo modelo.

Para Etchemendy (1999, p. 141), quando a abordagem Tarskiana é extensionalmente correta, como ocorre nos sistemas formais de primeira ordem clássicos, é uma questão de sorte. É por uma felicidade que o mundo tomado como base para a análise seja apropriado, considerando válidos todos e apenas os argumentos intuitivamente válidos.

Uma abordagem que considerasse a totalidade de mundos e de interpretações possíveis, considerando as características pré-teóricas da consequência lógica, não apresentaria os problemas das perspectivas modal e estrutural. A próxima proposta busca capturar as características fundamentais das abordagens modal e formal.

4 Abordagem modal/formal

A abordagem modal/formal, como o próprio nome indica, pretende considerar as características fundamentais das duas abordagens anteriores, visando eliminar os problemas oriundos de cada uma delas, quando sugeridas isoladamente.

De acordo com Hanson (1997, p. 37-8),

> Nesta visão a conclusão de um argumento é consequência lógica de suas premissas quando for *impossível* para o próprio argumento, ou para qualquer argumento com a mesma *forma*, possuir premissas verdadeiras e conclusão falsa. [itálicos nossos]

A noção de forma de um argumento utilizada aqui é a mesma exposta nos dois capítulos anteriores.

Construções desta perspectiva podem ser encontradas em Mendelson (1964), Shoenfield (1967), Quine (1967), Mates (1968) e Hanson (1997).

No que se segue, expomos, em linhas gerais, a proposta semântica modal/formal construída por Shoenfield (1967), em muitos aspectos semelhante à proposta de Tarski exposta na seção anterior.

Definição 3.4.1 (Estrutura para uma linguagem de primeira ordem)

Uma *estrutura* para uma linguagem de primeira ordem L^1, denotada por "$\mathcal{A}$", consiste dos seguintes elementos:

a: Um conjunto não-vazio $|\mathcal{A}|$ chamado *domínio* de $\mathcal{A}$. Os elementos de $|\mathcal{A}|$ são chamados os *indivíduos* de $\mathcal{A}$.

b: Para cada símbolo de função n-ária f de L^1, uma função n-ária $f_{\mathcal{A}}$ de $|\mathcal{A}|$ em $|\mathcal{A}|$ (em que $f_{\mathcal{A}}$ é um mapeamento do conjunto de n-uplas em $|\mathcal{A}|$ para $|\mathcal{A}|$). Em particular, para cada constante individual f de L^1, $f_{\mathcal{A}}$ é um indivíduo de $\mathcal{A}$.

c: Para cada letra de predicado P n-ário de L^1 distinta de =, um predicado n-ário $P_{\mathcal{A}}$ em $\mathcal{A}$ (em que $P_{\mathcal{A}}$ é um subconjunto do conjunto de n-uplas em $|\mathcal{A}|$).

Uma estrutura é uma "interpretação" ou uma atribuição de "significado" aos símbolos não-lógicos de uma linguagem, a partir dos quais são interpretadas todas as suas expressões. Juntamente com as próximas três definições, ela serve de base para as definições de verdade e consequência lógica.

Definição 3.4.2 (Linguagem estendida de uma linguagem de primeira ordem)

Uma *linguagem estendida de* L^1, denotada por "$L(\mathcal{A})$", é uma linguagem obtida a partir de L^1 acrescida de um nome de indivíduo (expressão que não pertence a L^1) para cada indivíduo de $\mathcal{A}$, de modo que indivíduos diferentes tenham nomes diferentes.

Nós utilizaremos o símbolo "η" como metavariável para os nomes de indivíduo de $L(\mathcal{A})$ e "g_j" para representar o j-ésimo nome de $L(\mathcal{A})$.

Definição 3.4.3 (Expressão livre de variável)

Uma expressão é *livre de variáveis* se não possuir variáveis.

Definição 3.4.4 (Interpretação de um termo livre de variável de L($\mathcal{A}$) em $\mathcal{A}$)

a: Se t é um nome de indivíduo de $L(\mathcal{A})$, então $\mathcal{A}(t)$ é o indivíduo de $\mathcal{A}$ nomeado por t.

b: Se t não é um nome de indivíduo de $L(\mathcal{A})$, então t é $ft_1...t_n$ e $\mathcal{A}(t)$ é $f_{\mathcal{A}}(\mathcal{A}(t_1)... \mathcal{A}(t_n))$.

Definiremos um valor de verdade para cada sentença φ em L($\mathcal{A}$). A definição, sugerida por Shoenfield (1967, p. 19), é por indução no comprimento de φ (número de símbolos lógicos presentes em uma fórmula φ).

Definição 3.4.5 (Valor de verdade de uma sentença)

O *valor de verdade* de uma sentença φ em $L(\mathcal{A})$, denotado por "$\mathcal{A}(\varphi)$", é assim definido:

1. Se φ é $t_0 = t_1$, então $\mathcal{A}(t_0 = t_1) = V$ sse $\mathcal{A}(t_0)$ é o mesmo que $\mathcal{A}(t_1)$, ou seja, $\mathcal{A}(t_0) =_{\mathcal{A}} \mathcal{A}(t_1)$.
2. Se φ é $Pt_0...t_n$, e P não é "=", então $\mathcal{A}(Pt_0...t_n) = V$ sse $P_{\mathcal{A}}(\mathcal{A}(t_0), ..., \mathcal{A}(t_n))$, ou seja, sse a n-upla $(\mathcal{A}(t_0), ..., \mathcal{A}(t_n)) \in P_{\mathcal{A}}$.
3. Se φ é $\neg\psi$, então $\mathcal{A}(\neg\psi) = V$ sse $\mathcal{A}(\psi) = F$.
4. Se φ é $(\psi \vee \gamma)$, então $\mathcal{A}((\psi \vee \gamma)) = V$ sse $\mathcal{A}(\psi) = V$ ou $\mathcal{A}(\gamma) = V$.
5. Se φ é $\exists x\, \psi$, então $\mathcal{A}(\exists x\, \psi) = V$ sse $\mathcal{A}(\psi_x[\eta]) = V$, para algum η em $L(\mathcal{A})$.

Definimos a seguir a noção de valoração para a Lógica Sentencial Clássica, *LSC* (Definição 2.2.16), a ser utilizada no próximo capítulo para comparar a semântica probabilística com a semântica veritativo-funcional para *LSC*.

Definição 3.4.6 (Valoração para *LSC* e valor de verdade de fórmulas)

a: Seja Format(*LSC*) o conjunto de fórmulas atômicas de *LSC*. Uma *valoração* para *LSC* é uma função $\mathcal{V}$: Format(*LSC*) $\rightarrow$ {V, F}.

b: Seja $\mathcal{V}$ uma valoração para *LSC*. O *valor de verdade* de uma fórmula $\varphi \in$ Form(*LSC*) segundo uma valoração $\mathcal{V}$ para *LSC*, denotado por "$\mathcal{V}(\varphi)$", é assim definido:

i: Se φ é atômica, então $\mathcal{V}(\varphi)$ é dado pela função $\mathcal{V}$.

ii: Se φ é $\neg\psi$ então $\mathcal{V}(\neg\psi) = V$ sse $\mathcal{V}(\psi) = F$.

iii: Se φ é $(\psi \vee \gamma)$ então $\mathcal{V}(\psi \vee \gamma) = V$ sse $\mathcal{V}(\psi) = V$ ou $\mathcal{V}(\gamma) = V$.

Dadas estas definições gerais, definimos a seguir a consequência lógica segundo a perspectiva em questão.

Definição 3.4.7 (Consequência lógica padrão)

Uma sentença φ é *consequência lógica padrão* de um conjunto Γ de sentenças, denotado por "$\Gamma \vDash \varphi$", sse, $\mathcal{Q}(\varphi) = V$ em toda estrutura $\mathcal{Q}$ tal que $\mathcal{Q}(\Gamma) = V$ (ou seja, $\mathcal{Q}(\varphi) = V$, para toda $\psi \in \Gamma$). (Em particular, φ é *consequência lógica padrão em LSC* de um conjunto Γ de sentenças sse, $\mathcal{V}(\varphi) = V$ em toda valoração $\mathcal{V}$ tal que $\mathcal{V}(\Gamma) = V$)

Utilizaremos a expressão "$\Gamma \nvDash \varphi$" para denotar que φ não é consequência lógica padrão de Γ. Denotaremos o conjunto de consequências lógicas padrão de Γ por $\vDash(\Gamma)$, ou seja: $\vDash(\Gamma) = \{\varphi \mid \Gamma \vDash \varphi\}$. Utilizaremos a expressão "CLP" como abreviação para "consequência lógica padrão".

Pode-se mostrar facilmente que CLP é uma consequência lógica Tarskiana, como era de se esperar.

A independência da definição de verdade dos símbolos não-lógicos de uma linguagem evita que CLP apresente o problema da complexidade da linguagem. Se o valor de verdade de uma sentença segundo uma estrutura $\mathcal{Q}$ não varia nas expansões ou contrações de uma linguagem, CLP também mantém-se inalterada nas suas expansões ou contrações.

Algumas semelhanças também podem ser observadas entre CLP e CLT. Ambas independem do significado dos símbolos não-lógicos das sentenças e são definidas com base em toda circunstância (estrutura ou linguagem). As duas são definidas a partir de um conjunto de símbolos lógicos. Em algum sentido, CLP substitui os símbolos não-lógicos das sentenças por metavariáveis, como definido em CLT.

O debate sobre as semelhanças entre CLP e CLT é bastante extenso. Etchemendy (1999) e Bays (2001), por exemplo, defendem a tese que Tarski adotou uma concepção de domínio fixo da consequência lógica. Para Bays (2001, p. 1702) tal adoção não origina os problemas apontados por Etchemendy (1999). Já Sher (1996), Gómez-Torrente (1996) e Ray (1996) argumentam a favor de Tarski ter sugerido uma concepção de modelo com domínios variados.

No artigo de 1933, Tarski parece pressupor a existência de universos distintos. Como a verdade é definida sempre para uma linguagem, a natureza dos indivíduos das sequências é estabelecida com base no objeto de discurso da linguagem. O próprio Tarski (1956c) constrói uma definição de verdade para uma teoria de conjuntos na qual os elementos das sequências são apenas conjuntos. Ele, porém, não se refere a domínios distintos, a subconjuntos do universo.

Já no artigo de 1936, Tarski parece pressupor um universo absoluto ao definir a noção de consequência lógica. Nele, estariam todos os objetos existentes, uma vez que CLT independe de qualquer linguagem ou interpretação dos símbolos não-lógicos. Sob esta ótica, concluiríamos que a definição de CLT não considera domínios distintos.

Outra diferença entre CLT e CLP estaria no mundo subjacente à definição de verdade. Em Tarski, ele parece ser o mundo *real*, como procuramos mostrar na seção anterior. Em CLP, o mundo subjacente é definido, à primeira vista, em cada estrutura. No restante desta seção analisamos a adequação de CLP considerando a diversidade de domínios e esta aparente independência do mundo *real*.

Etchemendy (1999, p. 74) afirma que uma das inovações da abordagem modal/formal com relação à abordagem original de Tarski refere-se à consideração de domínios distintos, o que implica em um novo tratamento dado aos quantificadores. Mas isso não salva esta abordagem de possuir alguns problemas, como procuramos mostrar a seguir.

Para evitar os problemas expostos nas abordagens modal e formal, além de não considerar lógicos aqueles símbolos que parecem não ter este caráter, como os nomes de indivíduos ou de predicados, os quantificadores também são tomados, em algum sentido, como símbolos variáveis. Isso possibilita variar a natureza dos objetos envolvidos nas sentenças e a quantidade de objetos a serem analisados, considerando as diferentes interpretações dos símbolos não-lógicos, nos diferentes mundos possíveis.

Etchemendy (1999, p. 65-9) apresenta duas versões para tornar variável um quantificador. Ambas propõem a sua interpretação e levam aos mesmos resultados. A interpretação de um quantificador é a atribuição de um conjunto de indivíduos da estrutura aplicados a ele. O quantificador universal "todo x" ("todo objeto", "todo elemento") pode ser interpretado como todo homem, todo cachorro, todo atual presidente do Brasil e o quantificador existencial "algum x" ("algum objeto", "algum elemento") como algum homem, algum cachorro, algum atual presidente do Brasil.

A interpretação dos quantificadores pode apresentar alguns problemas, como os ilustrados pelos seguintes argumentos:

Exemplo 3.4.10

a:	b:	c:
$\underline{\forall x\ \varphi}$	$\underline{Af^0{}_1}$	<u>Érico Veríssimo é escritor</u>
$\exists x\ \varphi$	$\exists x\ Ax$	Algum objeto é escritor

O argumento do Exemplo 3.4.10a, intuitivamente válido, poderia ser considerado inválido na perspectiva modal/formal caso os quantificadores universal e existencial fossem interpretados inversamente. Para evitar este problema, é necessária a restrição da interpretação dos quantificadores. Um quantificador universal não pode ser substituído por ou interpretado como quantificador existencial e vice-versa.

A restrição acima não resolve o problema do Exemplo 3.4.10b, considerado intuitivamente válido. Suponhamos que o Exemplo 3.4.10b seja "traduzido" para o português como em Exemplo 3.4.10c. Quando "$f^0{}_1$" for interpretado como Érico Veríssimo e "$\exists x$" como algum gato, considerando uma estrutura que espelhe o mundo *real*, a premissa de Exemplo 3.4.10b seria verdadeira e a sua conclusão seria falsa. Neste caso não teria havido equívoco na interpretação dos símbolos. As categorias semânticas teriam sido respeitadas. O problema consistiria na interpretação conjunta de "$Af^0{}_1$" e "$\exists x$". Segundo Etchemendy (1999, p. 65), para manter argumentos deste tipo válidos, a perspectiva modal/formal delimita restrições cruzadas de termos (*cross-term restrition*).

As restrições referem-se a proibições de interpretações conjuntas de termos. Um exemplo de tal restrição é aquela que obriga nomes de indivíduos e de predicados, quando relacionados como em Exemplo 3.4.10b, serem interpretados de tal modo que o indivíduo pertença à extensão do predicado.

Com a consideração de mundos possíveis distintos e a aplicação de domínios diferentes ao quantificador existencial, a abordagem modal/formal invalida, apropriadamente, os argumentos do Exemplo 3.3.8a' e do Exemplo 3.3.9.

O Exemplo 3.3.8a' é inválido devido ao conjunto das pessoas com menos de 35 anos e de presidente do Brasil não serem excludentes. Embora no mundo *real* todo presidente do Brasil tenha tido mais de 35 anos, é logicamente possível a existência de mundos em que pessoas com menos de 35 anos tenham sido presidentes. Na abordagem modal/formal isso poderia ser representado através de uma estrutura na qual algum indivíduo com menos de 35 anos pertencesse à extensão do predicado "presidente do Brasil".

Considerando as leis da lógica clássica aristotélica, em que $\vDash \forall x(x = x)$, qualquer *n*-instanciação do Exemplo 3.3.9 possui premissas verdadeiras e conclusão falsa em uma estrutura cujo domínio seja unitário, o que o torna inválido.

Ao impedir certas interpretações com a restrição cruzada de termos, a perspectiva modal/formal desconsidera a característica da necessidade. Tal exclusão poderia pressupor algum tipo de conhecimento prévio acerca dos objetos envolvidos nas sentenças, o que o faria desconsiderar a anterioridade.

Além dos problemas expostos acima, Etchemendy (1999, p. 113) procura mostrar que a abordagem modal/formal apresenta outras dificuldades. O seguinte esquema de argumento ilustra uma delas:

Exemplo 3.4.11

$\underline{\exists x\ (x = x \vee x \neq x)}$

$\neg\ \exists x_1 \ldots \exists x_n \prod^{n}_{i,j=1} (x_i \neq x_j)$, para $i < j$.

A conclusão do Exemplo 3.4.11 é equivalente a $\forall x_1 \ldots \forall x_n\ \Sigma^{n}_{i,j=1} (x_i = x_j)$, para $i < j$. Então:

Grau da instanciação	*Conclusão*
2	$\forall x_1 \forall x_2 (x_1 = x_2)$
3	$\forall x_1 \forall x_2 \forall x_3 (x_1 = x_2 \vee x_2 = x_3 \vee x_1 = x_3)$
4	$\forall x_1 \forall x_2 \forall x_3 \forall x_4 (x_1 = x_2 \vee x_2 = x_3 \vee x_3 = x_4 \vee x_1 = x_3 \vee x_1 = x_4 \vee x_2 = x_4)$

A conclusão do Exemplo 3.4.11 determina haver menos que *n* objetos no universo. Intuitivamente, qualquer *n*-instanciação deste argumento parece ser inválida, pois a conclusão não é garantida a partir da premissa. Para a perspectiva modal/formal ser adequada, deve haver uma estrutura tal que a premissa do Exemplo 3.4.11 é verdadeira e sua conclusão é falsa.

Para a conclusão de uma *n*-instanciação do Exemplo 3.4.11 ser falsa em uma estrutura, o seu domínio deve possuir pelo menos *n* indivíduos. Se concebermos intuitivamente que toda *n*-instanciação é inválida, para a proposta modal/formal ser adequada deve haver, para toda *n*-instanciação do Exemplo 3.4.11, uma estrutura em que a conclusão seja falsa e a premissa seja verdadeira. Se o universo for finito, contendo *k* objetos, toda *n*-especificação do Exemplo 3.4.11 tal que $n > k$ será válida e as demais serão inválidas. Logo, para a abordagem modal/formal ser adequada, deve pressupor que o universo seja infinito. Por outro lado, se concebermos intuitivamente que nem toda *n*-instanciação é válida, a adequação sempre dependeria da existência da quantidade de objetos no universo.

Uma das saídas para salvar a adequação da perspectiva modal/formal seria considerar a garantia da infinitude do universo uma característica lógica. Segundo Etchemendy (1999, p. 114), é isto o que se faz nessa perspectiva. Ela é construída com base na teoria de conjuntos do tipo ZF, nas quais o axioma da infinitude garante que o universo contém um número infinito de objetos.

Etchemendy (1999) argumenta que o axioma da infinitude expressa uma afirmação extra-lógica. Se fosse lógica, também o seria um axioma que afirmasse a existência de *n* objetos no universo. Mas isso tornaria válidos, indevidamente, uma série de *n*-instanciações do Exemplo 3.3.9 e Exemplo 3.4.11.

Afirmar o axioma da infinitude seria o mesmo que tentar adequar o mundo *real* escolhendo um indivíduo com menos de 35 anos presidente para garantir que argumentos como o do Exemplo 3.3.8a sejam inválidos, diz Etchemendy (1999, p. 115). Com essa afirmação, a abordagem modal/formal sofreria crítica semelhante à proposta Tarskiana original: pressuporia um mundo no qual a consequência lógica seria reduzida, o mundo subjacente ao das teorias de conjuntos em que figura o axioma da infinitude.

Uma solução para tais problemas poderia estar na readequação do conjunto de símbolos considerados genuinamente lógicos, desta vez eliminando dele o símbolo de predicado de igualdade "=". Isso tornaria inválidos os argumentos do Exemplo 3.3.9 e do Exemplo 3.4.11, tanto na abordagem Tarskiana quanto na modal/formal. Nessa circunstância, a conclusão de uma *n*-instanciação do Exemplo 3.3.9 afirma que o par ($\mathscr{A}(\tau_i)$, $\mathscr{A}(\tau_j)) \notin =_{\mathscr{A}}$, para algum par ($\mathscr{A}(\tau_i)$, $\mathscr{A}(\tau_j)$). Já a conclusão do Exemplo 3.4.11 é a negação da conclusão do Exemplo 3.3.9 e é formalmente equivalente a "$\forall x_1 \ldots \forall x_n \Sigma^{n}_{i,,j=1} (x_i=x_j)$, para $i < j$". Ela afirma que o predicado "$=_{\mathscr{A}}$" é diferente do vazio ϕ.

Para mostrar que o argumento do Exemplo 3.3.9 é inválido, basta apresentar uma estrutura em que o símbolo de predicado "=" é interpretado de tal modo que todo par de indivíduos do seu domínio pertence a $=_{\mathscr{A}}$. Etchemendy (1999, p. 117) cita os predicados *coexistir* e *ser idêntico ou não idêntico*, entendidos do modo usual, como exemplo em que isso ocorreria. Qualquer indivíduo de uma estrutura $\mathscr{A}$ coexiste com qualquer indivíduo de $\mathscr{A}$. Uma estrutura $\mathscr{A}$ que interpreta "=" como *coexistir*, torna falsa a conclusão do Exemplo 3.3.9 e verdadeira a sua premissa, para toda *n*-instanciação, independentemente do tamanho do universo ou do domínio.

Para mostrar que o argumento do Exemplo 3.4.11 é inválido, é suficiente exibir uma estrutura em que nenhum par de indivíduos pertença a $=_{\mathscr{A}}$. Uma estrutura com apenas um indivíduo na qual "=" é interpretado como *ser maior que*, entendido do modo usual, torna falsa a conclusão do Exemplo 3.4.11 e verdadeira a sua premissa, em toda *n*-instanciação do Exemplo 3.4.11.

De acordo com Etchemendy (1999, p. 117), a adequação na avaliação da validade do Exemplo 3.3.9 e do Exemplo 3.4.11 na proposta modal/formal independe de fatores extra-lógicos como o tamanho do universo. Porém, ao excluir o predicado de igualdade "=" do conjunto de símbolos genuinamente lógicos, CLP e CLT enfrentam problemas como os ilustrados pelos argumentos abaixo:

Exemplo 3.4.12

a:	b:
$A^1_1 f^0_1$	$\forall x\ \forall y\ \forall z$ (x é maior que y $\wedge$ y é maior que z $\rightarrow$ x é maior que z)
$\underline{a = f^0_2}$	$\underline{\forall x\ \neg(\text{x é maior que x})}$
$A^1_1\ f^0_2$	$\exists y\ \forall x\ \neg$(x é maior que y)

Quando "=" não é símbolo lógico, o argumento do Exemplo 3.4.12 é, erroneamente, considerado inválido, diz Etchemendy (1999, p. 166). Uma estrutura em que "=" é interpretado como *ser maior que* torna as premissas verdadeiras e a conclusão falsa.

De acordo com a conclusão do Exemplo 3.4.11b, *ser maior que* tem um elemento minimal. Já as suas premissas afirmam que este predicado é transitivo e irreflexivo. Como *ser maior que* é um "símbolo" não-lógico, a validade deste argumento depende de que, para qualquer predicado, se ele for transitivo e irreflexivo, então tem elemento minimal.

Pode-se interpretar um predicado atribuindo-lhe um conjunto de *n*-uplas ordenadas arbitrárias de indivíduos ou apenas como predicados genuínos. Segundo Etchemendy (1999, p. 118), no primeiro caso o Exemplo 3.4.12b é válido com relação a um domínio se, e somente se, esse domínio for finito, independentemente de haver restrições cruzadas de termos. Com isso, tem-se também que o Exemplo 3.4.12b é válido se, e somente se, o universo for finito. No segundo caso, o Exemplo 3.4.12b seria válido caso o universo fosse finito e poderia sê-lo caso o universo fosse infinito e homogêneo, por exemplo.

Em ambas as possibilidades de interpretação de um predicado, a abordagem modal/formal e a Tarskiana dependem do tamanho ou da constituição do universo para determinar a validade do Exemplo 3.4.12b. Mas, diz Etchemendy (1999, p. 120),

> ... qualquer um reconhece que a afirmação que não há um maior objeto não é uma consequência lógica do simples fato que a relação *maior que* é transitiva e irreflexiva. Nós o fazemos totalmente independente de nossas crenças ou hipóteses, quaisquer que sejam elas, sobre o tamanho real do universo. Sem dúvida nós simplesmente refletimos que não *poderia ter tido* nenhum maior objeto, havendo ou não de fato. Questões a respeito de o universo ser ou não finito são completamente irrelevantes para esta intuição fundamentalmente modal.

Etchemendy (1999, p. 121) procura mostrar que o problema das perspectivas formais da consequência lógica, especialmente a de Tarski e a modal/formal, não está na escolha apropriada de símbolos lógicos. Independente de como ele seja escolhido, há influências de fatores extra-lógicos como o tamanho do universo.

Quando aplicada aos sistemas formais lógicos de primeira ordem clássicos, a extensão de CLP é razoavelmente semelhante à extensão intuitiva de consequência lógica. Mas tal semelhança não se deve à sua correção ou a uma escolha particular de símbolos considerados genuinamente lógicos. A sua adequação pressupõe a infinitude do universo ou a sua não homogeneidade.

Como esta abordagem está construída com base em teorias de conjuntos do tipo ZF, em que há a garantia de que o universo é infinito e que seus elementos não são homogêneos, coincidentemente a extensão de CLP é equivalente à extensão do que se considera intuitivamente uma consequência lógica. Se o universo fosse distinto, talvez esta extensão não fosse a mesma.

Por depender de um mundo específico, CLP não possui a característica da necessidade. À primeira vista, ela parece satisfazê-la na medida em que considera mundos possíveis distintos. Segundo a crítica de Etchemendy (1999), o problema da abordagem modal/formal é que a possibilidade de um mundo é determinada segundo o mundo-base da teoria de conjuntos do tipo ZF. A consequência lógica é definida em função de um mundo particular. Mas, conclui Etchemendy (1999, p. 120), a relação de consequência lógica entre sentenças deve ser mantida independente da constituição do universo.

Ao tornar a consequência lógica dependente de um mundo específico, CLP deixa de possuir a característica da necessidade e da anterioridade. A abordagem modal/formal não considera todos os mundos e exige conhecimentos prévios sobre os objetos de discurso das sentenças envolvidas na análise.

Em suma, procuramos mostrar neste capítulo que, também sob o ponto de vista semântico, as abordagens aqui investigadas não satisfazem as três características da consequência lógica discutidas na Seção 1.2. Como visto, a validade de certos argumentos na abordagem modal depende do conteúdo das suas sentenças ou de conhecimentos prévios. Já a abordagem formal, em sua versão substitucional, permite que a validade de certos argumentos dependa da escolha dos termos lógicos de uma linguagem ou da complexidade da linguagem subjacente. Em sua versão interpretacional, a abordagem modal faz com que a validade de certos argumentos dependa, além da escolha dos símbolos lógicos de uma linguagem, também da análise do mundo *real* ou do tamanho do universo.

Por fim, na versão modal/formal a validade de certos argumentos também depende de fatores como a garantia da infinitude do universo. De um modo ou de outro, tais dependências de fatores extra-lógicos faz com que cada uma destas abordagens, que são usuais e aceitas quase que universalmente no âmbito da Lógica, deixem de satisfazer alguma das características da consequência lógica.

Nestes três primeiros capítulos deste trabalho procuramos fazer uma análise geral de alguns estudos clássicos sobre a consequência lógica. Procuramos apresentar as características centrais das propostas estudadas, investigando se elas satisfazem as principais características da consequência lógica. Em termos semânticos, a concepção usual base para a definição de consequência lógica é a de verdade. No restante deste trabalho proporemos uma concepção de consequência lógica baseada na noção de quantidade de informação. Introduziremos uma definição de consequência lógica informacional assumindo como ponto de partida a Lógica Sentencial clássica, definida no segundo capítulo. Investigaremos qual a lógica subjacente, dentre os sistemas expostos anteriormente, à consequência lógica informacional e se ela satisfaz as características de necessidade, formalidade e anterioridade.

Capítulo 4
O conceito de informação

1 Apresentação

Neste capítulo iniciamos a segunda parte do trabalho, na qual construímos nossa proposta de uma análise informacional da consequência lógica. Apresentamos a noção quantitativa de informação, adotada em áreas como a Teoria Matemática da Comunicação, a ser utilizada como base para a nossa proposta de uma inferência informacional. Na próxima seção expomos a noção de informação subjacente a esta perspectiva matemática da informação. Inicialmente apesentamos a noção de modelo de comunicação unidirecional, com vistas a definir uma fonte de informações, a ser tomada como referência para, no sexto capítulo, estabelecermos o valor informacional de uma fórmula que, por sua vez, será utilizada na definição da inferência informacional. Devido aos nossos objetivos neste trabalho, não trataremos do aspecto semântico da informação, mas apenas do seu aspecto quantitativo. Na terceira seção apresentamos a definição de quantidade de informação e finalizamos o capítulo com algumas críticas a esta perspectiva.

2 A noção quantitativa de informação

Do ponto de vista da Teoria Matemática da Comunicação (doravante, TMC), a comunicação é vista como um processo de transmissão de informações, a exemplo do que ocorre em uma ligação telefônica, um envio de e-mail ou em uma consulta em terminais eletrônicos. De acordo com Shannon e Weaver (1949, p. 03),

> A comunicação é qualquer procedimento pelo qual uma mente afeta outra mente. Além da fala escrita e oral, a comunicação envolve música, artes pictóricas, teatro, balé e, de fato, todo comportamento humano. Em algumas situações pode ser desejável usar uma definição mais ampla de comunicação. Tal definição envolveria procedimentos por meio dos quais um mecanismo (por exemplo um equipamento automático para rastear um aeroplano e computar suas prováveis posições futuras) afeta outro mecanismo (por exemplo um míssil guiado perseguindo este aeroplano).

A transmissão de informações pressupõe alguns componentes, como uma fonte, um canal e um destino. Esses três elementos são entidades do mundo, diz Blahut (1988, p. 3).

A fonte e o receptor podem ser células nervosas, computadores, seres humanos, a natureza; o canal pode ser o espaço ou um conjunto de fios.

Para que a informação seja transmitida pela fonte ao destino através do canal, geralmente é necessária a construção de dois componentes. Um deles para adequar as mensagens à natureza do canal para que elas possam ser transmitidas e outro para readequar estas mensagens para que o destino possa recebê-las. Estes dois componentes são denominados, respectivamente, transmissor e receptor.

A figura abaixo ilustra o modelo de comunicação unidirecional, no qual as informações são transmitidas apenas em uma direção: da fonte para o destino.

Figura 4.2.1

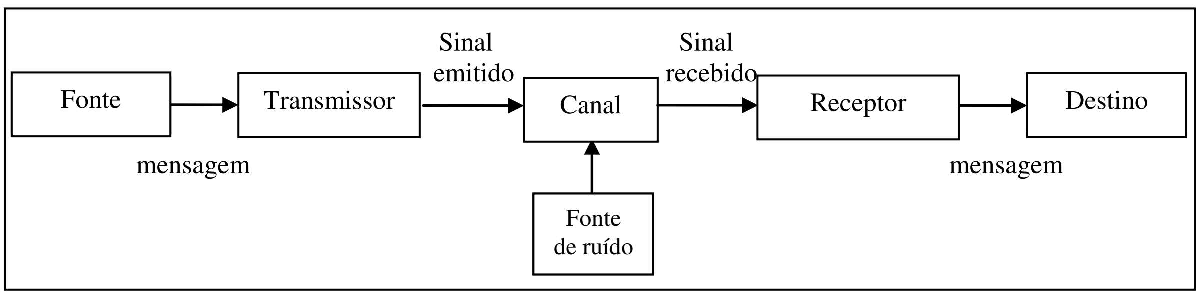


Figura 4.2.1
Modelo de comunicação unidirecional (Shannon & Weaver, 1949, p. 07)

No exemplo da ligação telefônica, o falante é a fonte de informações e o ouvinte o destino. As palavras emitidas pela fonte são transformadas em sinais pelo transmissor e transmitidas através de fios ou via satélite para o telefone do ouvinte, onde são novamente transformadas em sons pelo receptor a fim de serem recebidas pelo destino.

Em um sentido geral, uma fonte pode ser entendida como um processo gerador de informações. Os elementos de uma *fonte no mundo* são os eventos possíveis de serem realizados, isto é, fenômenos em inter-relação, delimitados e gerados pela própria fonte. Um evento pode ser pensado como um acontecimento, um comportamento, uma realização de algo, atual ou possível, dependendo da circunstância.

Exemplos de fontes são os lances de moeda ou de dados, cujos eventos são, respectivamente, cara e coroa e cada um dos seis lados da face do dado; os jogos de loteria, cujos eventos são os números a serem sorteados; o clima, cujos eventos são chuva, calor, umidade etc.; a Língua Portuguesa, cujos eventos podem ser considerados as suas letras ou as suas palavras.

Em termos teóricos, uma fonte pode ser pensada como um processo gerador de mensagens, as portadoras de informação, concebidas como um conjunto de signos organizados de acordo com determinadas regras adotadas. Para Wiener (1970, p. 33) "... uma mensagem é uma sequência discreta ou contínua de elementos mensuráveis distribuídos no tempo (o que os estatísticos chamam série temporal)." Pode ser uma sequência de símbolos ou de palavras escritas ou faladas de uma linguagem, a temperatura registrada por um termômetro contínuo ou o toque de uma campainha.

Assim, cada conjunto de eventos, parte constituinte de uma fonte no mundo, pode ser associado a um conjunto de mensagens, parte constituinte de uma fonte no sentido teórico. Dizer que uma mensagem foi gerada ou escolhida significa que ocorreu o seu evento correspondente no mundo.

A rigor, quem ocorre ou possui probabilidade de ocorrência são os eventos e não as mensagens de uma fonte. No restante deste trabalho, no entanto, diremos, por questões práticas, que as mensagens ocorrem ou possuem probabilidade de ocorrência. Ao dizer que a mensagem de uma fonte ocorreu, ou possui certa probabilidade de ocorrência, deve-se compreender que o seu evento correspondente em uma dada circunstância no mundo ocorreu ou possui a probabilidade estabelecida.

As mensagens podem ser digitais ou analógicas. As digitais, ou discretas, são constituídas por elementos com duração e tamanho delimitados, como nas sentenças escritas de uma linguagem. Nelas, é possível distinguir, enumerar, classificar e identificar com precisão os seus elementos. Já as mensagens analógicas, ou contínuas, se caracterizam por não apresentar separação nítida entre seus componentes, como nas sentenças faladas ou nos velocímetros que registram a velocidade através de ponteiros em vez de números.

Uma fonte discreta é aquela que manipula (escolhe, delimita e gera) apenas mensagens discretas. Tais fontes geram suas mensagens símbolo por símbolo, como na telegrafia ou na computação digital, em que os elementos das mensagens podem ser representados por sequências finitas de dígitos binários "0" e "1".

Em geral, as mensagens são selecionadas com base na probabilidade de ocorrência das mesmas. Quanto maior a probabilidade de ocorrência, maior é a chance da mensagem ser selecionada pela fonte.

A produção de uma sequência de símbolos de acordo com certas probabilidades é chamada *processo estocástico*. Para Shannon (1949, p. 40), uma fonte discreta pode ser representada por um processo estocástico, ou seja, como um espaço de probabilidades. Por outro lado, um processo estocástico que produz uma sequência discreta de símbolos escolhidos a partir de um conjunto finito pode ser considerado uma fonte discreta.

Quando, em um processo estocástico, as probabilidades dependem da ocorrência de mensagens anteriores, tem-se um processo ou uma *cadeia de Markov*. De acordo com Epstein (1986, p. 59), uma cadeia de Markov pode ser definida como um processo probabilístico no qual o desenvolvimento futuro depende estatisticamente do estado presente. No Português, por exemplo, a probabilidade de ocorrência da letra "e" é muito maior do que a probabilidade da letra "t". A chance de "e" ser escolhida é muito maior do que "t". Já a probabilidade de ocorrência de "t", dada a ocorrência de "m" é nula, o que não acontece com a letra "b".

Dentre as cadeias de Markov estão os *processos ergóticos*, caracterizados pela existência de uma regularidade estatística segura. Como suas propriedades não mudam com o tempo, qualquer amostra razoavelmente grande de suas mensagens tende a ser representativa da sequência geral de ocorrência de suas mensagens. Segundo Wiener (1970, p. 33), a previsão do futuro de uma mensagem faz-se sobre uma espécie de operador sobre o seu passado, seja ele realizado por um esquema de computação matemática ou por um aparelho mecânico ou elétrico.

Fontes ergóticas são aquelas em que toda sequência produzida tem as mesmas propriedades estatísticas que qualquer outra. As suas propriedades não se alteram com o tempo. Descobertas as probabilidades de ocorrência dos símbolos, pode-se prever, para qualquer momento, a probabilidade de ocorrência daquele símbolo.

No próximo capítulo trataremos da noção de situação, que corresponderemos a fontes discretas ergóticas. Elas serão tomadas como base para a proposta da semântica probabilística e da noção de consequência lógica informacional.

No que se segue propomos uma concepção de fonte discreta ergótica com base na proposta de Shannon (1949).

Definição 4.2.2 (Fonte discreta ergótica)

Sejam $A_F = \{x_1, ..., x_n\}$ e $P_F = \{p(x_1), ..., p(x_n)\}$, em que $p(x_i) \in \mathbb{Q}$, $0 \leq p(x_i) \leq 1$, $\Sigma^n_{i=1} p(x_i) = 1$. Uma *fonte discreta ergótica F* é tal que $F = \{(x_1, p(x_1)), ..., (x_n, p(x_n))\}$.

Na definição acima, "A_F" é denominado o espaço da fonte *F*, x_i é a *i*-ésima mensagem da fonte *F*, P_F é o conjunto dos valores probabilidade das mensagens de *F*, de tal modo que $p(x_i)$ é o valor probabilidade da *i*-ésima mensagem de *F*.

A fonte constituída pelas mensagens correspondentes aos eventos de um lance de moedas, por exemplo, em que a mensagem "cara" é representada pelo dígito 0 e "coroa" pelo dígito 1, seria assim descrita: {(0,1/2); (1,1/2)}. Já a fonte constituída pelo lance de um dado seria: {(1,1/6); (2,1/6); (3,1/6); (4,1/6); (5,1/6); (6,1/6)}.

Na perspectiva da TMC, a eficiência da comunicação consiste em transmitir, de modo adequado, as informações selecionadas na fonte para o destino. Os problemas de comunicação, dizem Shannon & Weaver (1949, p. 04), podem ser estudados sob três enfoques, representados pela solução dos seguintes problemas:

a: Problema técnico: quão acuradamente os símbolos de comunicação podem ser transmitidos.

b: Problema semântico: quão precisamente os símbolos transmitidos carregam o significado desejado.

c: Problema da efetividade: quão efetivamente o significado recebido afeta a conduta de modo desejado.

O *problema técnico* consiste na análise da exatidão da transferência da mensagem selecionada pela fonte. Sob esse prisma, avalia-se em que medida os símbolos que chegaram ao destino são, de fato, os que partiram da fonte. Investigam-se ainda as causas das possíveis falhas no processo de comunicação, como o ruído (perturbações que podem alterar as mensagens originais) e os meios para evitá-los, eliminá-los e corrigi-los, como faz a TMC. A informação é analisada sob um ponto de vista quantitativo e sintático. Neste aspecto, a eficiência da comunicação é medida pela reprodução exata ou aproximada no destino dos símbolos selecionados na fonte.

O *problema semântico*, por sua vez, refere-se à precisão do significado dos sinais. Nessa perspectiva busca-se analisar em que medida o destino capturou o significado das mensagens emitidas pela fonte. A informação é concebida sob um ponto de vista qualitativo e semântico.

A eficiência da comunicação semântica pressupõe a eficiência da comunicação sintática. Em geral, para que o significado original da mensagem emitida pela fonte seja transmitido e captado pelo destino, é essencial que os símbolos constituintes da mensagem cheguem ao destino sem modificações significativas, ou seja, com o mínimo possível de ruído. No problema semântico a análise é centrada naquilo que as mensagens se referem ou sobre o seu conteúdo. Aquilo que elas querem dizer. A questão a ser investigada consiste na análise da qualidade da mensagem, o seu significado ou a sua referência.

Por fim, o *problema da efetividade* consiste na análise da realização dos comandos subjacentes na mensagem emitida pela fonte. Uma vez recebida a mensagem e captado o seu significado pelo destino, investiga-se em quais condições a sua conduta é satisfatória. A eficiência da comunicação é definida a partir da adequação da atividade realizada pelo destino com os pressupostos da mensagem na fonte.

A solução do problema da efetividade pressupõe a solução dos outros dois problemas anteriores. Se o destino não captura o significado da mensagem ou a recebe de modo alterado, pode não capturar o pedido implícito emitido pela fonte.

Para Shannon e Weaver (1949, p. 31), os aspectos semânticos e pragmáticos da informação são irrelevantes para o problema técnico, ou de engenharia. Porém, aspectos de engenharia não são irrelevantes para os aspectos semânticos ou de efetividade.

Muitas críticas foram dirigidas ao modelo de comunicação unidirecional, especialmente com relação à sua adequação à explicação da comunicação e ação humanas. Netto (2001), por exemplo, analisa modelos de comunicação em que a seleção de informações não está centrada na fonte, mas também no destino. Para Netto (2001, p. 200), o modelo unidirecional não é adequado para explicar a dinâmica do funcionamento da ação humana não-mecânica. Para Le Coadic (1996), um bom modelo de comunicação social não pode ser direcional. Para ele, o processo de comunicação humana é circular, sem a presença de fonte e destino. Todos os participantes do processo comunicativo informam e são informados ao mesmo tempo. Nesse processo, tanto a informação quanto a ação poderiam emergir de forma auto-organizada, no sentido explicitado por Debrun (1996*a*; 1996*b*).

Embora as críticas esboçadas ao modelo unidirecional sejam muito expressivas, o objetivo de Shannon (1949), e de grande parte dos pesquisadores da TMC, não era o de explicar a natureza do processo comunicativo humano ou processos comunicativos que envolvem aspectos semânticos ou pragmáticos. Eles estavam interessados, particularmente, com o problema da eficiência técnica, ou seja, da transmissão exata de um conjunto de símbolos de um ponto para outro.

Um dos recursos utilizados para medir a eficiência da comunicação foi estabelecer uma definição quantitativa da informação. Antes de apresentarmos essa definição da medida informacional, apresentamos a seguir a noção de informação subjacente à proposta de Shannon.

Dois dos pioneiros no estudo da quantificação, do armazenamento e da transmissão da informação foram Nyquist e Hartley (1928). Eles descrevem a quantidade de informação presente em uma fonte de acordo com o seu número de mensagens possíveis. Shannon (1948) aprimorou essa ideia e estabeleceu a base para a TMC. Ele inclui novos fatores à proposta de Nyquist e Hartley, como o efeito do ruído no canal, a economia possível na transmissão de informações e a possibilidade de mensagens possuírem quantidade de informações distintas.

Na perspectiva em questão, diz Pignatari (1968, p. 45), só pode haver informação onde há dúvida e dúvida implica na existência de alternativas – donde escolha, seleção, discriminação.

Para Hartley (1928), a informação em uma mensagem é medida pela liberdade de escolha que alguém tem ao selecioná-la, baseada em uma fonte. De acordo com Shannon & Weaver (1949, p. 8), a informação relaciona-se não ao que você *realmente* diz, mas ao que *poderia* dizer. É uma medida da liberdade de escolha quando se seleciona uma mensagem.

Em um lance não viciado de moeda, por exemplo, há duas possibilidades igualmente prováveis de escolha: cara ou coroa. Já em um lance não viciado de dados, há seis possibilidades. A liberdade de escolha no primeiro caso é menor que no segundo. No caso dos dados, poderíamos dizer muito mais coisas do que poderia ser dito no caso da moeda. Por isso, intuitivamente, pode-se perceber, sob a ótica em questão, que a quantidade de informação presente no jogo de dados é maior do que a do lance de moeda.

Para Hershberger (1955) a informação pode ser definida como uma medida da redução de incerteza. Está relacionada à imprevisibilidade em uma mensagem ou em uma fonte, trazendo à tona um elemento ausente antes da sua ocorrência. No lance de moeda a redução de incerteza é menor do que no jogo de dado. A ocorrência de um evento em uma fonte como a do exemplo do lance de moeda elimina apenas uma alternativa, enquanto no lance de dados são eliminadas cinco alternativas equiprováveis.

A informação pressupõe a possibilidade de ocorrência de mais de uma mensagem em uma fonte. Se ela admite apenas a ocorrência de uma mensagem, não há liberdade de escolha, não há incerteza a ser reduzida. Sabe-se *a priori* qual mensagem ocorreria, o que impede a geração de novidade ou redução de incerteza. Uma fonte caracterizada pelo jogo de um dado no qual todos os lados tivessem o mesmo número gravado seria não informativa, porque poderíamos saber de antemão o resultado do lance.

Quanto maior a liberdade de escolha, a redução de incerteza, em uma fonte, mais informativa ela é. A informação atinge seu valor máximo quando todas as mensagens de uma fonte tiverem a mesma chance de serem escolhidas. Já o valor mínimo ocorre quando apenas uma delas puder ocorrer.

A informação também costuma ser associada à noção de ordem. Na TMC o termo "ordem" é utilizado para se referir à noção que se caracteriza pela estabilidade, regularidade, arranjo e pressupõe previsibilidade, regularidade. A desordem, por sua vez, é caracterizada pela aleatoriedade, pelo acaso, pela randomicidade. Em um texto recente sobre conceitos básicos de sistêmica, Bresciani & D'Ottaviano (2000) definem os conceitos de ordem e organização. Nesta proposta, que não será discutida neste trabalho, a noção de organização é correspondente à noção de ordem na TMC.

Na TMC a ordem em uma fonte é definida a partir da distribuição da probabilidade de seus eventos ou mensagens. Uma fonte totalmente desordenada é aquela cujos eventos ou mensagens possuem a mesma probabilidade de ocorrência. Já a ordem máxima ocorre quando um evento possui probabilidade absoluta de ocorrência. Quanto mais díspares forem as probabilidades de ocorrência dos eventos, mais ordenada é a fonte.

A ordem é geralmente associada à noção de entropia, definida como a medida da aleatoriedade de uma fonte. Aplicada à noção de informação, a entropia é definida com base na probabilidade de ocorrência das mensagens de uma fonte. Na concepção de Shannon & Weaver (1949, p. 12), a entropia é a medida da incerteza de uma variável randômica. A entropia é a medida da desordem. Quanto mais desordenada uma fonte, maior a sua quantidade de entropia. A entropia está em proporção inversa à ordem.

O aumento na entropia significa um aumento da liberdade de escolha ou da redução da incerteza. Quanto mais desordenada for uma fonte, maior será sua quantidade média de informação. Por isso, dizem Shannon e Weaver (1949, p. 15), a informação e a entropia estão na mesma proporção.

Quanto mais semelhantes forem as probabilidades de ocorrência das mensagens de uma fonte, maior é a sua desordem. A desordem está na mesma proporção da aleatoriedade ou randomicidade. Quando, em uma fonte, as mensagens acontecem aleatoriamente, a previsibilidade é sacrificada. A ordem pressupõe que algumas coisas aconteçam mais vezes e outras menos, ou seja, depende da disparidade da probabilidade de ocorrência das mensagens.

Nesse processo, uma fonte é totalmente desordenada, isto é, possui o máximo de entropia, quando todas as suas mensagens são equiprováveis. Em fontes com essa característica, quanto maior a sua quantidade de mensagens possíveis, mais ela é desordenada. Por outro lado, por maior que seja o número de mensagens possíveis em uma fonte, se uma delas tiver uma probabilidade muito elevada de ocorrência, digamos, quase absoluta, a desordem é baixa.

Dado este panorama geral da noção de informação subjacente à proposta de Shannon, na próxima seção expomos uma definição da quantidade de informação presente em uma mensagem e em uma fonte.

3 Uma definição da medida informacional

A definição sugerida por Shannon (1949) para a medida da liberdade de escolha em uma fonte pode ser ilustrada através de um método de divisão das mensagens, como exposto por Edwards (1971), por exemplo. Esse procedimento consiste em sucessivas divisões do número de alternativas possíveis em dois grupos com o mesmo número de elementos, escolhendo um desses grupos em cada divisão com base na resposta à pergunta: “Escolho o primeiro grupo?” O número de respostas “sim” ou “não” necessárias até obter dois grupos constituídos por apenas um elemento será o valor da liberdade de escolha. Claramente, esse método funciona apenas quando o número de mensagens da fonte é tal que na última pergunta tem-se um par de elementos. Além disso, ele é adequado unicamente quando as mensagens possuem a mesma probabilidade de ocorrência.

De acordo com esse método, uma fonte com oito mensagens equiprováveis precisaria de três respostas à pergunta em questão, conforme ilustrado na figura abaixo que, como as próximas figuras, é adaptada da exposição de Edwards (1971):

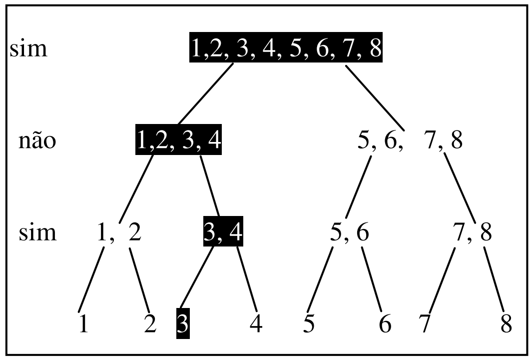


Figura 4.3.1
Método de divisão das mensagens de uma fonte com oito mensagens possíveis equiprováveis

A seleção ou a geração da mensagem número 3, portanto, demandou três escolhas sucessivas. Esse método permite estabelecer a seguinte relação:

Número de mensagens (n)	divisão por 2	número de decisões (H)
1	2	0
2	2	1
4	2	2
8	2	3
16	2	4

Figura 4.3.2
Relação entre o número de mensagens de uma fonte e o de decisões para a escolha de um delas.

Este método parece se adequar à noção intuitiva de informação exposta na seção anterior. Quanto maior for o número de elementos do espaço da fonte, maior será a liberdade de escolha, definido pelo número destas escolhas, a ser associado à quantidade de informação.

De acordo com a Figura 4.3.2 acima, estabelecemos, por combinatória, a seguinte relação matemática entre o número de mensagens, denotado por "*n*", e o número de perguntas, denotado por "*I*":

$$n = 2^{I}$$

De acordo com a definição de função logarítmica, se B = A^x então $\log_A B = x$. A partir disso, Shannon (1949) estabelece uma definição da quantidade de informação em uma fonte com uma quantidade de mensagens equiprováveis.

Definição 4.3.3 (Quantidade de informação em fonte com mensagens equiprováveis) Seja *F* uma fonte de informações com *n* mensagens equiprováveis. A *quantidade de informação de F*, denotada por "I_F", é definida pela seguinte equação:

$$I_F = \log_2 n.$$

Fontes com quatro, oito ou dezesseis mensagens equiprováveis conteriam, respectivamente, quantidade de informação igual a 2, 3 e 4. Assim, a fonte composta pelo jogo de moeda é menos informativa do que a fonte do jogo de dados.

Muitas fontes de informação, no entanto, possuem mensagens com probabilidade de ocorrência distinta. Algumas ocorrem mais vezes do que outras, como no caso de um lance de moeda viciado, por exemplo.

Na perspectiva de Shannon (1949), a rigor, a informação é atribuída à fonte e não às suas mensagens individuais. Segundo Weaver (1949, p. 9), embora algumas vezes seja conveniente atribuir informação às mensagens individuais (como se faz com o significado), o objetivo é definir uma quantidade de informação para uma fonte como um todo. No entanto, a definição da quantidade de informação em uma fonte de informações cujas mensagens não são equiprováveis depende de uma espécie de quantificação da informação presente em cada mensagem.

Quando as mensagens não são equiprováveis, o método da divisão em grupos com o mesmo número de mensagens para determinar a quantidade de informação deixa de ser apropriado. Como a liberdade de escolha diminui nestes casos, dada a existência de uma espécie de preponderância de algumas mensagens, o número de perguntas para a escolha de uma mensagem também deve diminuir. Quando as mensagens são equiprováveis, a divisão do número de mensagens ao meio é viável porque cada metade fica com cinquenta por cento de probabilidade de ocorrer, o que nem sempre acontece quando elas não são equiprováveis.

Um novo método para medir a quantidade de informação nesses casos consiste em considerar a probabilidade de ocorrência das mensagens em uma fonte de tal modo que possibilite a divisão em grupos igualmente prováveis de serem escolhidos.

Para ilustramos esse novo método, imaginemos uma fonte com quatro mensagens em que, de cada oito ocorrências de mensagens, a mensagem "1" ocorre quatro vezes, a mensagem "2" ocorre duas vezes e as mensagens "3" e "4" ocorrem uma vez cada. A figura abaixo, adaptada do texto de Edwards (1971), ilustra como adequar o caso de fontes, cujas mensagens possuem probabilidades distintas, ao método anterior.

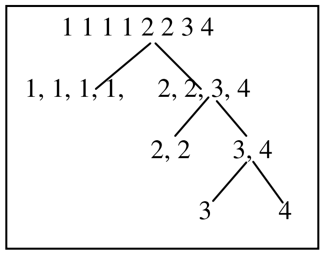


Figura 4.3.4
Método de divisão das mensagens de uma fonte com quatro mensagens possíveis

Com esta representação podemos proceder como no caso das mensagens equiprováveis, dividindo o grupo ao meio. Mas, como ilustra a figura acima, a seleção da primeira mensagem exigiria apenas uma escolha, da segunda exigiria duas e das duas últimas exigiria três escolhas. O esquema abaixo ilustra essa relação.

Mensagem (i)	divisão	n. de escolhas (I_i)	probabilidade (p_i)
1	2	1	½
2	2	2	¼
3	2	3	1/8
4	2	3	1/8

Figura 4.3.5
Relação entre o número de escolhas para a seleção de uma mensagem com a sua probabilidade de ocorrência.

No esquema acima encontramos a seguinte relação matemática entre as três últimas colunas:

$$p_i = 2^{-I_i}.$$

De acordo com a definição da função logarítmica, Shannon (1949) estabelece a quantidade de informação presente em uma mensagem de uma fonte, conforme a próxima definição.

Definição 4.3.6 (Quantidade de informação de uma mensagem)

Seja x_i uma mensagem de uma fonte *F* e p_i a probabilidade de ocorrência de *i*. A quantidade de informação em x_i, denotada por "I_i", é definida pela seguinte equação:

$$I_i = - \log_2 p_i.$$

Quando $p_i = 0$, definimos que $- \log_2 0 = 0$, ou seja, $I_i = 0$.

A quantidade de informação presente em cada uma das mensagens enunciadas na figura 4.3.4 acima é assim definida:

$$I_1 = - \log_2 ½ = 1 \qquad I_2 = - \log_2 ¼ = 2$$
$$I_3 = - \log_2 1/8 = 3 \qquad I_4 = - \log_2 1/8 = 3$$

Pode ser facilmente mostrado que, em fontes com mensagens equiprováveis, a quantidade de informação de qualquer das suas mensagens é igual à quantidade de informação da própria fonte. Nestes casos, a liberdade de escolha de uma mensagem coincide com a liberdade de escolha da própria fonte.

Além disso, deriva-se das definições acima que a quantidade de informação de uma mensagem é sempre positiva. Não há informação negativa na perspectiva de Shannon (1949).

A determinação da quantidade média de informação de fontes em que as mensagens não são equiprováveis não pode basear-se na sua quantidade de mensagens. A quantidade de informação presente em uma mensagem pode ser determinada pelo número de decisões binárias necessárias para poder selecioná-la. A quantidade média de informação presente em uma fonte pode ser determinada pelo número médio dessas decisões para cada uma de suas mensagens.

No exemplo da Figura 4.3.4 acima, observamos que a quantidade de vezes que a primeira mensagem é selecionada é muito maior do que o número de vezes que as demais mensagens. Como a sua probabilidade de ocorrência é maior, a sua quantidade de informação deve ser menor. O número de decisões para a sua escolha também deve ser menor, o que representa que a liberdade de escolha nessa fonte diminui.

Intuitivamente, o número médio de decisões depende da probabilidade de ocorrência de cada mensagem associado ao número de decisões exigido para ela ser selecionada. Neste método, verifica-se que quanto mais uma mensagem tem chance de ser escolhida, menos decisões são necessárias para a sua seleção. No caso do exemplo da Figura 4.3.4, o número médio de decisões da fonte em questão, denotado por "H", é assim calculado:

$$H = (½ \text{ x } 1) + (¼ \text{ x } 2) + (1/8 \text{ x } 3) + (1/8 \text{ x } 3) = 1, 75.$$

O resultado da equação acima mostra que o grau de liberdade de escolha diminuiu quando comparado à fonte em que as mensagens eram equiprováveis. Houve uma economia no número médio de escolhas.

Para calcular H procedemos inicialmente multiplicando a probabilidade de ocorrência de cada mensagem pela sua quantidade de informação

$$I_i\text{: } p_i \times \log_2 p_i.$$

Com isso, obtemos a quantidade de informação em uma fonte qualquer, conforme a próxima definição.

Definição 4.3.7 (Quantidade de informação presente em uma fonte)
Sejam F uma fonte de informações com n mensagens e I_i a quantidade de informação presente na mensagem x_i de F. A quantidade de informação presente em F, denotada por "H_F", é definida pela seguinte equação:

$$H_F = \Sigma_i\, p_i \times I_i$$

Abaixo consideramos três fontes e calculamos a sua quantidade de informação.

Exemplo 4.3.8
$F_1 = \{(a, ½); (b, ½)\}$.
$F_2 = \{(a, ¼); (b, ¼); (c, ¼); (d, ¼)\}$.
$F_3 = \{(a, ½); (b, ¼); (c, 1/8); (d, 1/8)\}$.

H_{F1} = - {(½ -1) + (½ -1)} = 1.

H_{F2} = - {(¼ -2) + (¼ -2) + (¼ -2) + (¼ -2) = 2.

H_{F3} = - {(½ -1) + (¼ -2) + (1/8 -3) + (1/8 -3) = 1,75.

As duas primeiras fontes do exemplo acima apresentam mensagens equiprováveis. Por conter mais mensagens, a segunda fonte apresenta uma quantidade de informação maior que a primeira. A terceira fonte, por sua vez, embora tenha o mesmo número de mensagens do que a segunda, apresenta uma quantidade de informação menor. Isso acontece porque as mensagens na terceira fonte não são equiprováveis.

Pode-se mostrar, como o faz Roman (1992), por exemplo, que uma fonte cujas mensagens são equiprováveis constitui um caso particular da definição acima.

A próxima figura ilustra a variação da informação de uma fonte H com duas mensagens, denotadas por "x_1" e "x_2". A probabilidade de ocorrência de x_1 é denotada por "$p(x_1)$", de modo que $p(x_2)$ = "1 - $p(x_1)$".

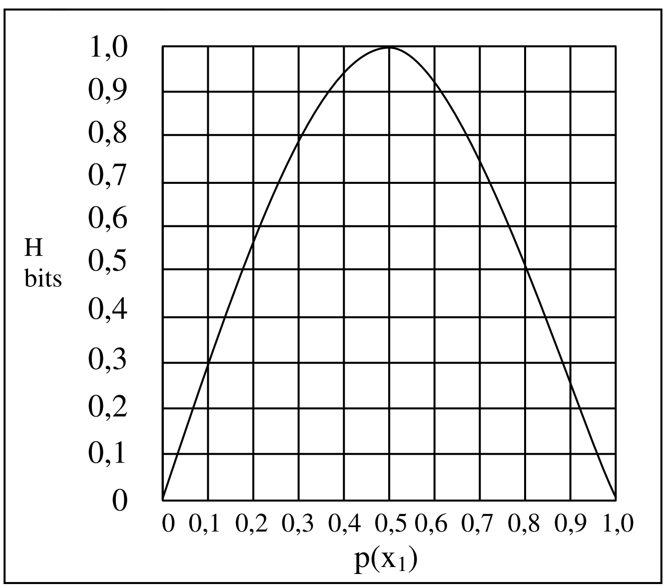


Figura 4.3.9

Variação na informação de uma fonte com duas mensagens (Shannon & Weaver, 1949, p. 50)

O gráfico explicita, dentre outras coisas, que a informação é máxima quando $p(x_1) = p(x_2)$ = ½. Já a informação é mínima quando uma das mensagens possui a probabilidade máxima.

Embora esta perspectiva quantitativa da informação tenha a sua utilidade, ela sofre inúmeras críticas.

Para Stonier (1990), há um paradoxo na caracterização de informação de Shannon, que pode ser ilustrado através do seguinte exemplo: imaginemos uma biblioteca cujos livros estão distribuídos por assunto, autor, palavras-chave. Nela, é fácil encontrar uma obra solicitada. Diríamos que tal biblioteca é altamente informativa (considerando a distribuição física dos livros e não o conteúdo destes), pois podemos saber onde estão os seus livros com grande facilidade. Mas, segundo a perspectiva de Shannon, ela conteria uma pequena quantidade de informação, dado o grau elevado de ordem. Assim, quanto mais informativa for uma fonte, menos informativa ela parece ser.

Acreditamos que o paradoxo acima se origina devido à ambiguidade da noção de informação, o que o torna um pseudo-paradoxo. Por um lado, sob a perspectiva de Shannon, a informação está associada à desordem. Por outro lado, conforme áreas como a ciência da informação, a informação é aquilo que uma fonte diz efetivamente sobre algum estado de coisas; é um conhecimento inscrito ou gravado na forma escrita ou falada, conforme Yves (1996).

Para Stonier (1990, p. 07) a informação não é nem matéria nem energia. Mas, assim como a matéria e a energia, ela existe no mundo físico. A sua existência independe dela ser percebida ou entendida, ter um significado ou ser interpretada, diz Stonier (1990, p. 22). A informação presente no DNA ou nos símbolos gravados em uma pedra existe independente de ela ser compreendida por algum captador específico. Se os signos vierem a ser decifrados, então o DNA ou a pedra, além de conter (*contain*) informação, também a carrega/exprime (*convey*).

Na visão do pensador em questão, informação e ordem estão na mesma proporção. Quanto mais ordenado for um sistema, mais informação ele carrega. A informação organiza o espaço e o tempo. Ela é definida por Stonier (1990, p. 26) como a capacidade para organizar um sistema ou para mantê-lo em um estado ordenado. A ordem é a manifestação da informação interagindo com matéria e energia. No exemplo da distribuição física dos livros de uma biblioteca, quanto mais ordenada ela for, mais informações conterá. Se mudarmos um livro de lugar, haverá uma mudança informacional.

Se a mudança gera desordem, por exemplo, retirando um livro do devido lugar, há uma perda informacional. Se, ao contrário, pusermos o livro no seu lugar correto (um livro de filosofia na estante de filosofia, não na de física), então teremos uma biblioteca mais ordenada. Embora a informação dependa da ordem material da biblioteca, ela não pode ser confundida com a própria matéria.

A ordem de um sistema reflete o arranjo das suas partes constituintes. Stonier (1990, p. 33) propõe uma relação inversa entre informação e desordem: quanto mais desordenado for um sistema, menor é o seu conteúdo informacional.

A entropia negativa é uma medida de ordem, conclui Stonier (1990, p. 38). A informação é uma função exponencial inversa da entropia, entendida como uma medida da desordem, que se contrapõe à ordem de um sistema, em especial, de uma fonte. Quanto maior a entropia em um sistema, menor a informação presente nele. Um sistema é menos ordenado na medida em que seus elementos tendem a ser distribuídos randomicamente, afirma Stonier (1990, p. 37). Um cristal, cuja entropia é baixa, apresenta uma quantidade informacional quase absoluta. Uma biblioteca cujos livros são distribuídos com base em algum padrão contém mais informação do que aquela cujas obras estão jogadas aleatoriamente nas estantes.

Dretske (1981) também direciona algumas críticas à concepção da informação sugerida por Shannon. Ele reconhece o valor dessa proposta e a utiliza para propor uma concepção semântica de informação.

Na visão de Dretske (1981, p. 40), "Uma teoria genuína da informação seria uma teoria sobre o conteúdo de nossas mensagens e não sobre a forma pela qual este conteúdo é incorporado." Não se pode confundir o sinal que carrega uma informação com a própria informação. Seria o mesmo que confundir o balde que carrega a água ou a quantidade de água no balde com a própria água, ilustra o pensador. A TMC mede a quantidade de informação transmitida de uma fonte a um destino através de um canal, mas não diz o que está sendo transmitido.

Para Dretske (1981, p. 44), a informação encontrada em um sinal (signo que representa uma mensagem na fonte) é o que ele é capaz de dizer *verdadeiramente* sobre algum estado de coisas no mundo. A informação é aquele artigo capaz de produzir conhecimento. Não há informação quando o que está sendo transmitido não corresponde à realidade do objeto na fonte. Quando tal correspondência existe, a informação carrega um *significado natural* sobre os eventos no meio ambiente. Um conjunto de nuvens pretas, por exemplo, significa a possibilidade de chuva para um receptor (humano ou não) atento às regularidades da natureza.

Além da correspondência entre mensagem e mundo, o conteúdo informacional de uma mensagem depende também do conjunto de informações acumuladas no destino das informações, como ilustra Dretske (1981, p 65): em uma mesa há quatro conchas enfileiradas e sob uma delas está escondido um amendoim. Dois indivíduos, A e B, criam um jogo cujo vencedor é quem descobre primeiro em qual delas está escondido o alimento. Suponhamos que o indivíduo A, mas não o B, saiba que o objeto não está nas duas primeiras conchas. O ato de levantar a terceira concha carrega o conteúdo informacional para o indivíduo A, mas não para B, da localização exata do amendoim.

Dretske (1981) utiliza a noção shannoniana de informação (no tocante ao seu aspecto objetivo e quantitativo) para apresentar uma caracterização de conhecimento. Dretske (1981, p. 86) define que quando há uma quantidade de informação positiva associada ao caso de s ser F, K conhece que s é F é o mesmo que: a crença de K de que s é F é causada (ou causalmente sustentada) pela informação de que s é F. Nessa caracterização, “K” é um sistema capaz de conhecer (ter crenças, manipular informações); “s” é um elemento demonstrativo ou indexical que se refere a algum elemento de uma fonte; “F” é um predicado pertencente a uma sentença. Ser “causalmente sustentado” é entendido como sinônimo da existência de uma relação entre a informação na fonte e a crença gerada por ela no destino.

Para melhor ilustrar os termos da caracterização de Dretske, apresentamos o seguinte exemplo: um sujeito (*K*) observa uma mesa em uma sala e acredita que essa mesa (*s*) é quadrangular (*F*). Se tal crença for sustentada essencialmente pela informação de que a mesa (*s*) é quadrangular (*F*), e por outros conhecimentos, observações empíricas ou definições, então o sujeito (*K*) conhece/sabe que a mesa é quadrangular (*s é F*).

Devlin (1991) também realiza uma investigação da parte semântica da informação e concorda que ela é uma entidade existente no mundo. Ele utiliza a lógica para determinar o conteúdo informacional de uma fonte. Para tratar da informação, diz Devlin (1991, p. 10) “... uma “lógica” baseada na verdade (tal como a lógica clássica) não é apropriada; o que se exige é uma “lógica” baseada na informação.”

Para Devlin (1991, p. 39), a informação é algo que resulta da combinação de uma restrição, de um recorte (*constraint*) e de uma representação de eventos e situações no mundo. Uma restrição é algo que liga vários tipos de situações, seja por meio de leis naturais, convenções, regras analíticas, linguísticas. A sentença “fumaça implica fogo” expressa uma restrição do tipo lei natural. A relação de dependência estabelecida entre o toque de uma campainha e a presença de alguém à porta também é uma restrição. Uma restrição relaciona uma fonte a um conjunto de fontes; é por meio dele que se determina a quais fontes uma dada fonte pode ser relacionada.

A representação torna perceptível a primeira fonte da restrição, como a fumaça ou o ruído provocado pelo toque na campainha. Se o receptor estiver sintonizado tanto com a restrição (conhecer a relação entre o tocar na campainha e a existência de alguém que a toque) quanto com a representação (for capaz de ouvir o som típico da campainha), é capaz de receber a informação (há alguém à porta) que resulta da restrição e da representação.

Nas últimas páginas procuramos mostrar a inexistência de um consenso mínimo a respeito da caracterização da informação. Como ressalta Devlin (1991, p. 03),

> ... estamos em situação quase similar àquela do homem da era do ferro que, apesar de manipular e viver cercado por instrumentos de ferro, não dispunha de instrumentos conceituais apropriados para explicar a natureza química ou física desse elemento.

Vivemos na Era da Informação, mas ainda não temos uma noção minimamente consensual do que ela seja. Talvez a sua compreensão pressuponha, inclusive, um arcabouço teórico distinto do que dispomos no momento.

Em suma, procuramos expor, neste capítulo, a perspectiva informacional desenvolvida por Shannon e adotada em parte na TMC. Na perspectiva quantitativa, a informação está relacionada à liberdade de escolha ao se selecionar uma mensagem, ao que se poderia dizer, e não ao que efetivamente se diz, em uma dada circunstância. Tal liberdade está associada ao valor probabilidade de cada mensagem, definida a partir do logaritmo na base dois deste valor. Um dos pressupostos básicos desta perspectiva é a inexistência de informação negativa. Isso explica o sinal negativo na equação que define o valor probabilidade de uma mensagem, garantindo que o valor informacional de um evento seja sempre positivo. Outro pressuposto básico é o de que cada mensagem deve possuir um valor determinado em uma dada fonte. Isto justifica a arbitrariedade em definir como nulo o valor informacional de mensagens com valor probabilidade zero, justificado também pelo pressuposto de que estas mensagens nada informam sobre o “mundo”. Voltaremos a este assunto no Capítulo 6, mais especificamente nos comentários sobre a Definição 6.2.1. Na perspectiva de Shannon, informação e entropia estão na mesma proporção; quanto mais desordenada for uma fonte, maior será a sua quantidade de informação. No final deste capítulo comentamos algumas críticas a esta perspectiva quantitativa, tanto no que diz respeito à sua relação com a entropia, quanto no tocante à própria noção de informação subjacente a esta proposta. Apesar das críticas, nos próximos dois capítulos utilizaremos a perspectiva quantitativa como ponto de partida para uma definição de uma inferência informacional.

Capítulo 5
Uma Semântica Probabilística para a Lógica Sentencial Clássica

1 Apresentação

Neste capítulo apresentamos uma semântica para a Lógica Sentencial Clássica, *LSC*, baseada na noção de probabilidade. Na próxima seção tratamos de alguns elementos de uma teoria axiomática de probabilidades usual, expondo algumas das suas definições e resultados básicos a serem utilizados posteriormente em nosso trabalho. Na terceira seção construímos uma semântica probabilística para *LSC*. Estabelecemos uma relação entre as fórmulas de *LSC* com os eventos de um experimento aleatório, a partir da qual definiremos um valor probabilidade para cada fórmula de *LSC*. Mostramos algumas características e resultados elementares desta semântica aplicados às fórmulas, enfatizando especialmente aqueles referentes à implicação material. Na última seção introduzimos uma definição de fórmulas probabilisticamente válidas e probabilisticamente equivalentes e mostramos alguns resultados gerais da semântica probabilística, comparando-a à semântica veritativo-funcional usual.

2 Elementos de uma teoria axiomática de probabilidades

Assumiremos como base para nossa proposta de uma concepção informacional de consequência lógica a noção de probabilidade. Nesta seção apresentamos algumas definições e resultados básicos da Teoria de Probabilidades usual, como expostos por lógicos como Carnap (1962). Eles serão utilizados posteriormente na construção da semântica probabilística e na definição de consequência lógica informacional.

Fazemos uso da probabilidade nos casos em que dois ou mais resultados diferentes podem ocorrer. Isto torna o resultado não previsível (ou não determinado), no sentido de que não há como determinar previamente qual resultado ocorrerá em um dado momento. A *Teoria da Probabilidade*, doravante denotada por “$\mathbb{P}$”, estuda experimentos aleatórios, cuja noção será utilizada, juntamente com outras noções básicas dessa teoria, dentre eles os de espaço amostral e evento, como ponto de partida para nossa proposta.

Utilizaremos como base para $\mathbb{P}$ a Teoria de Conjuntos de Zermelo-Fraenkel (ZF) com a Teoria Aritmética Elementar usual. Assim, a linguagem (alfabeto e definições) e os teoremas de ZF serão também considerados elementos de $\mathbb{P}$. Os únicos símbolos próprios do alfabeto de $\mathbb{P}$ são as ocorrências de "A_i", para $0 \leq i \leq n$ e $i \in \mathbb{N}$, ou seja, A_0, A_1, A_2, etc. O símbolo "A" representa o conceito primitivo de $\mathbb{P}$ denominado *resultado* ou *acontecimento*, utilizado para a definição de experimento aleatório, exposta a seguir.

Definição 5.2.1 (Experimento aleatório)

Um *experimento* ou *fenômeno aleatório*, denotado por "Σ", é aquele que, repetido diversas vezes, apresenta diferentes resultados ou acontecimentos, denominados *resultados de Σ* ou *acontecimentos de Σ* e denotados por "$A_i(\Sigma)$".

Como exemplos de experimentos aleatórios podemos citar o lançamento de uma moeda ou de um dado, a retirada de uma carta em um baralho, a realização de um campeonato esportivo. Seus resultados poderiam ser, respectivamente, a queda da moeda voltada com a face cara ou coroa para cima, a queda do dado com um dos números de um a seis no lado voltado para cima, a retirada de cada uma das cartas do baralho, os campeões do campeonato em cada edição.

As próximas definições são baseadas nesta noção de experimento aleatório.

Definição 5.2.2 (Espaço amostral de Σ)

a: O *espaço amostral* de um experimento aleatório Σ, denotado por "$U(\Sigma)$", é o conjunto de todos os resultados possíveis de Σ.

b: O *número de elementos* do espaço amostral, denotado por "$n(U(\Sigma))$", em que $n(U(\Sigma)) > 0$ e finito, é a quantidade de elementos de "$U(\Sigma)$".

c: Um espaço amostral é *equiprovável* quando todos os seus elementos possuem a mesma chance de ocorrer.

Em outras palavras, para um experimento aleatório Σ cujos resultados possíveis são $A_1(\Sigma), ..., A_n(\Sigma)$, tem-se que $U(\Sigma) = \{A_1(\Sigma), ..., A_n(\Sigma)\}$.

Para o experimento Σ_1, que consiste no lançamento de uma moeda não-viciada, tem-se que $U(\Sigma_1) = \{C, K\}$ e $n(U(\Sigma_1)) = 2$, ou seja, são dois os resultados possíveis, cara, "C", ou, nos termos da teoria, “$A_1(\Sigma_1)$”, “ e coroa, “K”, ou “$A_2(\Sigma_1)$”. Para o experimento Σ_2, que consiste no lançamento de um dado não viciado, tem-se que $U(\Sigma_2) = \{1, 2, 3, 4, 5, 6\}$ e $n(U(\Sigma_2)) = 6$. Para o experimento Σ_3, que consiste no lançamento de duas moedas, tem-se que $U(\Sigma_3) = \{(C, C), (C, K), (K, C), (K, K)\}$ e $n(U(\Sigma_3)) = 4$.

Os espaços amostrais dos exemplos acima são equiprováveis, o que não ocorre no lançamento de uma moeda viciada, em que, por exemplo, de cada quatro lances, três saem cara. No entanto, neste experimento, denominemos Σ_4, podemos considerar $U(\Sigma_4) = \{C_1, C_2, C_3, K\}$ e $n(U(\Sigma_4)) = 4$, o que resultaria em um espaço amostral equiprovável. O mesmo ocorre com o exemplo do campeonato esportivo, em que o espaço amostral é constituído das equipes campeãs em cada edição do campeonato; o espaço poderia não ser equiprovável, dada a possibilidade de algumas equipes serem campeãs com maior frequência do que outras.

Doravante, quando não gerar ambiguidade, eliminaremos as referências a Σ entre parênteses das notações. Assim, em vez de $A(\Sigma)$, $U(\Sigma)$ ou $n(U(\Sigma))$, escreveremos apenas A, U ou n(U), respectivamente.

Definição 5.2.3 (Evento de um experimento aleatório)

a: Um *evento* de um experimento aleatório Σ, denotado por “$E(\Sigma)$”, é qualquer subconjunto do espaço amostral $U(\Sigma)$, ou seja, $E(\Sigma) \subseteq U(\Sigma)$.

b: O *número de elementos* de um evento $E(\Sigma)$, denotado por “$n(E(\Sigma))$”, é a quantidade de elementos de $U(\Sigma)$ pertencentes a $E(\Sigma)$.

A definição acima diz que um evento de Σ é um conjunto constituído unicamente por resultados possíveis de Σ. Claramente, um mesmo evento pode fazer parte de diferentes experimentos aleatórios. Por exemplo, {1} pode ser um evento do jogo de dados ou do jogo de loteria.

Os acontecimentos de Σ, em geral, não são eventos de Σ, por não serem subconjuntos de $U(\Sigma)$. O espaço amostral $U = \{\{1\}, 1\}$ configura um exemplo de um experimento Σ no qual {1} seria tanto um acontecimento quanto um evento de Σ.

Essa diferença entre acontecimento e evento pode ser explicada com base no âmbito em que cada um deles está inserido. O acontecimento diz respeito ao que efetivamente ocorre no experimento aleatório. Já o evento é uma entidade teórica correspondente aos acontecimentos de um “mundo”.

Mesmo nos casos em que um elemento é tanto um acontecimento quanto um evento de um experimento, como no exemplo acima, a sua natureza enquanto acontecimento é distinta da sua natureza enquanto evento. Conforme explicitaremos na Definição 5.2.4c abaixo, um evento pode ser vazio, no sentido da Teoria de Conjuntos, o que não é permitido a um acontecimento. Sequer faz sentido falar de acontecimento vazio, dado que esta noção não se aplica a acontecimentos, mas a conjuntos. Poderíamos dizer, no entanto, que um resultado vazio seria aquele que não ocorre. Mas então ele já não seria um resultado.

Expostos os primeiros conceitos de $\mathbb{P}$, apresentamos, a seguir, algumas definições oriundas deles. Iniciamos com uma classificação de diferentes tipos de eventos.

Definição 5.2.4 (Classificação de eventos de um experimento aleatório)

a: Um *evento elementar E de Σ* é aquele tal que $n(E(\Sigma)) = 1$.

b: O *evento certo E de Σ* é aquele tal que $n(E(\Sigma)) = n(U(\Sigma))$.

c: O *evento impossível E de Σ*, denotado por “$\varnothing$”, é aquele tal que $n(E(\Sigma)) = 0$.

d: Um *evento contingente E de Σ* é aquele tal que $n(\varnothing) < n(E(\Sigma)) < n(U(\Sigma))$.

e: O *evento E_i de Σ complementar de E(Σ)*, denotado por “$\bar{E}(\Sigma)$”, é definido por $\bar{E}(\Sigma) = \{A \in U(\Sigma) \mid A \notin E(\Sigma)\}$.

f: O *evento E de Σ união de $E_i(\Sigma)$ e $E_j(\Sigma)$*, denotado por “$(E_i(\Sigma) \cup E_j(\Sigma))$”, é definido por $(E_i(\Sigma) \cup E_j(\Sigma)) = \{A \in U(\Sigma) \mid A \in E_i(\Sigma) \text{ ou } A_\Sigma \in E_j(\Sigma)\}$.

g: O *evento E de Σ interseção de $E_i(\Sigma)$ e $E_j(\Sigma)$*, denotado por “$(E_i(\Sigma) \cap E_j(\Sigma))$”, é definido por $(E_i(\Sigma) \cap E_j(\Sigma)) = \{A_\Sigma \mid A_\Sigma \in E_{i\Sigma} \text{ e } A_\Sigma \in E_{j\Sigma}\}$.

h: Dois eventos $E_i(\Sigma)$ e $E_j(\Sigma)$, são *mutuamente exclusivos* sse $n(E_i(\Sigma) \cap E_j(\Sigma)) = 0$.

Pode-se facilmente mostrar, com base na Teoria de Conjuntos ZF, a existência de exatamente um evento certo e de exatamente um evento impossível em cada experimento aleatório, como indicado pela definição acima. Do mesmo modo, existe um e apenas um complemento para um evento e exatamente uma união e uma interseção para dois eventos quaisquer. Já a quantidade de eventos elementares e contingentes depende da complexidade do experimento aleatório.

Todo evento elementar é constituído por um único elemento do espaço amostral e o evento certo de Σ é o próprio espaço amostral de Σ. O evento impossível de Σ, por ser o conjunto vazio, apesar de não poder ser um acontecimento, é evento de todo experimento. Se um evento é elementar ou impossível em um experimento, será, respectivamente, elementar ou impossível em todo experimento do qual for evento. No entanto, ele pode ser certo em um experimento e contingente em outro e vice-versa.

No exemplo a seguir apresentamos alguns casos de experimentos aleatórios com parte de seus elementos constituintes.

Exemplo 5.2.5

Σ	**U(Σ)**	**E(Σ)**	**n(E(Σ))**
$\mathbf{\Sigma_1}$ (Lance de moeda)	{C, K}	$E_1(\mathbf{\Sigma_1})$: {C} (Sair cara) $E_2(\mathbf{\Sigma_1})$: {K} (Sair coroa)	1 1
$\mathbf{\Sigma_2}$ (Lance de moeda viciada)	{C_1, C_2, C_3, K}	$E_1(\mathbf{\Sigma_2})$: {C_1, C_2, C_3} (Sair cara) $E_2(\mathbf{\Sigma_2})$: {K} (Sair coroa)	3 1
$\mathbf{\Sigma_3}$ (Lance de dado)	{1, 2, 3, 4, 5, 6}	$E_1(\mathbf{\Sigma_3})$: {2,4,6} (Sair par) $E_2(\mathbf{\Sigma_3})$: ∅ (Sair cara)	3 0

Quadro 5.2.5: Exemplos de experimentos aleatórios

No exemplo acima está proposto um “modelo” para cada experimento aleatório, seu espaço amostral e alguns de seus eventos, associando-lhes a entidades de um “mundo”. Na terceira coluna, entre parênteses, determinamos, para cada evento, o seu nome na Língua Portuguesa, procurando expressar os resultados que o constituem.

A noção de experimento aleatório pode ser comparada com a noção de fonte discreta ergótica de informações, tal como na Definição 4.2.2. Como neste tipo de fonte, exigiremos que os elementos pertencentes a um experimento aleatório devem estar prévia e precisamente definidos. Além disso, todo espaço amostral considerado, além de finito, deverá ser equiprovável e os valores probabilidade dos eventos devem ser dados e fixos. Isso nos permitirá determinar com precisão a existência da relação de consequência lógica entre fórmulas.

A próxima definição trata da probabilidade dos eventos em um dado experimento aleatório.

Definição 5.2.6 (Probabilidade de um evento)

A *probabilidade de ocorrência* de um evento E no experimento aleatório Σ com um espaço amostral equiprovável U(Σ), denotada por "p(E(Σ))", é o valor numérico definido pela seguinte equação:

$$p(E(\Sigma)) = \frac{n(E(\Sigma))}{n(U(\Sigma))}.$$

A *função probabilidade*, p, leva eventos de um experimento aleatório a valores entre 0 e 1, ou seja, p: $E(\Sigma) \rightarrow [0,1] \subseteq \mathbb{Q}$. O *valor probabilidade* do evento E(Σ) é dado por p(E(Σ)).

Considerando o Exemplo 5.2.5, apresentamos abaixo um quadro de possibilidades para os eventos lá sugeridos.

Exemplo 5.2.7

E	$E_1(\Sigma_1)$	$E_2(\Sigma_1)$	$E_1(\Sigma_2)$	$E_2(\Sigma_2)$	$E_1(\Sigma_3)$	$E_2(\Sigma_3)$	$E_1(\Sigma_3) \cup E_2(\Sigma_3)$	$E_1(\Sigma_2) \cap E_2(\Sigma_2)$	$\bar{E}_1(\Sigma_2)$
p(E)	½	½	¾	¼	½	0	½	0	¼

Quadro 5.2.7: Probabilidade de eventos do Exemplo 5.2.5

Constituídos os elementos básicos da Linguagem de ℙ, a seguir estabelecemos os seus axiomas e alguns de seus resultados.

Os axiomas para ℙ são os seguintes:

(Axℙ1): $p(E) \geq 0$, para todo $E \subseteq U$.

(Axℙ2): $p(E_i \cup E_j) = p(E_i) + p(E_j) - p(E_i \cap E_j)$.

(Axℙ3): $p(E \cup \bar{E}) = 1$.

Esses três axiomas, que correspondem aos axiomas da Teoria de Probabilidades usual, fazem sentido para nossos objetivos tanto intuitivamente quanto para as definições previamente expostas.

O (Axℙ1) determina que o valor probabilidade mais baixo de um evento E é zero, o que é intuitivamente óbvio. Não faria sentido dizer que um evento tem probabilidade negativa de ocorrer. Esse axioma também está de acordo com $\mathbb{P}$. Com base na Definição 5.2.6, $p(E) = n(E)/n(U)$. Mas, pela Definição 5.2.3, $n(E) \geq 0$, dado que o menor subconjunto de um conjunto é o conjunto vazio, e, pela Definição 5.2.2, $n(U) > 0$ e finito. Como o denominador é finito, e ambos os valores $n(E)$ e $n(U)$ são números naturais, então o valor probabilidade de todo evento é um número racional. Quando $n(E) = 0$, $p(E) = 0$ e quando $n(E) > 0$, $p(E) > 0$.

O (Ax$\mathbb{P}$1) estabelece que a probabilidade do evento união de dois eventos é a soma da probabilidade de ambos, subtraída da sua interseção. Isso também parece adequado, tanto intuitiva quanto teoricamente.

Intuitivamente, a união de dois eventos consiste em reunir todos os elementos pertencentes a cada um deles. O número da união é o resultado da soma do número de cada um deles. No entanto, se houver elementos que pertencem aos dois eventos, eles serão somados duas vezes, gerando um resultado maior do que a quantidade efetiva de elementos dos dois conjuntos. Para o número de elementos da união ser adequado, necessitamos subtrair o número de elementos pertencentes a ambos eventos, de modo que cada um seja somado uma única vez. Com isso, permitimos uma definição adequada do valor probabilidade da união.

Teoricamente, com base na Teoria de Conjuntos, sabe-se que $n(E_i \cup E_j) = n(E_i) + n(E_j) - n(E_i \cap E_j)$. De acordo com a Definição 5.2.6, $p(E_i \cup E_j) = n(E_i \cup E_j)/n(U)$. Aplicando resultados da Aritmética usual, temos que $p(E_i \cup E_j) = n(E_i)/n(U) + n(E_j)/n(U) - n(E_i \cap E_j)/n(U)$. Outra vez, pela Definição 5.2.6, $p(E_i \cup E_j) = p(E_i) + p(E_j) - p(E_i \cap E_j)$. Assim, (Ax$\mathbb{P}$2) concorda tanto com a intuição tanto com as definições de $\mathbb{P}$.

O (Ax$\mathbb{P}$3) determina que a probabilidade da união de um evento com seu complementar é a unidade. Isso também é óbvio. Pela Definição 5.2.4e–f, tal união compreende todos os elementos do espaço amostral. Assim, $n(E \cup \bar{E}) = n(U)$. Pela Definição 5.2.6, tem-se que $p(E \cup \bar{E}) = n(E \cup \bar{E})/n(U) = n(U)/n(U) = 1$.

Dados os axiomas de $\mathbb{P}$, expomos a seguir alguns de seus resultados. No primeiro deles demonstramos duas associações entre o número de elementos de dois eventos de um mesmo experimento e a probabilidade deles nesse experimento. Os símbolos metalinguísticos "$\Leftrightarrow$" e "$\Rightarrow$" representam, respectivamente, as expressões "Se, e somente se" e "se, então", interpretados da maneira usual na Língua Portuguesa.

Teorema 5.2.8: Sejam E_i e E_j dois eventos de um experimento aleatório Σ.

a: $n(E_i) = n(E_j) \Leftrightarrow p(E_i) = p(E_j)$.

b: $E_i \subseteq E_j \Rightarrow p(E_i) \leq p(E_j)$.

Demonstração

a: $n(E_i) = n(E_j) \Leftrightarrow n(E_i)/n(U) = n(E_j)/n(U) \Leftrightarrow_{Def5.2.6}\Leftrightarrow p(E_i) = p(E_j)$.

b: $(E_i \subseteq E_j) \Rightarrow n(E_i) \leq n(E_j) \Rightarrow n(E_i)/n(U) \leq n(E_j)/n(U) \Rightarrow_{Def5.2.6}\Rightarrow p(E_i) \leq p(E_j)$. ■

É trivial, como um caso particular do Teorema 5.2.8b, que $E_i = E_j \Rightarrow p(E_i) = p(E_j)$. No entanto, a recíproca deste resultado não pode ser mostrada, ou seja, $p(E_i) = p(E_j) \not\Rightarrow E_i = E_j$. No Exemplo 5.2.5, $p(E_1(\Sigma_1)) = p(E_2(\Sigma_1))$, mas $E_1(\Sigma_1) \neq E_2(\Sigma_1)$. Esse resultado só é garantido nos casos em que $E_i = U$ ou $E_i = \varnothing$.

Em geral, a recíproca do Teorema 5.2.8b não pode ser demonstrada em $\mathbb{P}$, como ilustra o Exemplo 5.2.5: $p(E_2(\Sigma_2)) = 1/4 \leq p(E_1(\Sigma_2)) = 3/4$, mas $E_2(\Sigma_2) = \{K\} \not\subset E_1(\Sigma_2) = \{C_1, C_2, C_3\}$.

Abaixo demonstramos algumas propriedades elementares satisfeitas pela probabilidade.

Teorema 5.2.9

a: $p(E) \leq 1$, para todo $E \subseteq U$.

b: $p(U) = 1$.

c: $p(\varnothing) = 0$.

d: $n(E_i \cap E_j) = 0 \Rightarrow p(E_i \cup E_j) = p(E_i) + p(E_j)$.

e: $p(E \cap \bar{E}) = 0$.

f: $p(\bar{E}) = 1 - p(E)$.

g: $\sum^{m}_{i=1} p(E_i) = 1$, para $E_i = \{A_i\}$, para $1 \leq i \leq m$, com $U = \{A_1, ..., A_m\}$.

Demonstração

a: Caso 1: $E = U \Rightarrow n(E) = n(U) \Rightarrow_{Def5.2.4}\Rightarrow p(E) = 1$.

Caso 2: $E \subset U \Rightarrow n(E) < n(U) \Rightarrow p(E) = n(E)/n(U) < 1$.

b: $p(U) =_{Def5.24}= n(U)/n(U) = 1$.

c: $p(\varnothing) =_{Def5.2.6}= n(\varnothing)/n(U) =_{Def5.2.4c}= 0/n(U) = 0$.

d: $n(E_i \cap E_j) = 0 \Rightarrow_{Def5.2.4c}\Rightarrow E_i \cap E_j = \varnothing$. Mas $p(E_i \cup E_j) =_{(Ax\mathbb{P}2)}= p(E_i) + p(E_j) - p(E_i \cap E_j) = p(E_i) + p(E_j) - p(\varnothing) =_{Teo5.2.9c}= p(E_i) + p(E_j) - 0 = p(E_i) + p(E_j)$.

e: $E \cap \bar{E} =_{Def5.2.4}= \varnothing \Rightarrow p(E \cap \bar{E}) = p(\varnothing) =_{Teo5.2.9c}= 0$.

f: $n(E \cap \bar{E}) =_{Def5.2.4}= 0 \Rightarrow_{Teo5.2.9d}\Rightarrow p(E \cup \bar{E}) = p(E) + p(\bar{E}) =_{(Ax\mathbb{P}3)}= 1 \Rightarrow p(\bar{E}) = 1 - p(E)$.

g: Seja $E_i = \{A_i\}$, para $1 \leq i \leq m$, com $U = \{A_1, ..., A_m\}$. Pela Definição 5.2.4a, $n(E_i) = 1$. Além disso, $n(U) = m$. Pela Definição 5.2.6, $p(E_i) = 1/m$. Assim, $\sum^{n}_{i=1} p(E_i) = p(E_1) + ... + p(E_n) = 1/m + ... + 1/m$ (*m* vezes) = 1. ■

Do Teorema 5.2.9a e do (Ax$\mathbb{P}$1) temos que o valor probabilidade de um evento é um número racional entre zero e um, ou seja, $0 \leq p(E) \leq 1$.

A segunda e terceira propriedades acima determinam os eventos com maior e menor valor probabilidade em um experimento aleatório. Como era de se esperar, eles são os eventos certo e impossível.

A quarta propriedade é um caso especial de (Ax$\mathbb{P}$2) e define a probabilidade da união de dois eventos mutuamente exclusivos como a soma de suas probabilidades. Já a quinta propriedade estabelece que a probabilidade da interseção de um evento com seu complementar é nula. É o oposto do que ocorre com a união desses dois eventos, estabelecido pelo (Ax$\mathbb{P}$3).

A sexta propriedade define a probabilidade do complementar de um evento. Por fim, a última propriedade mostra que a somatória das probabilidades de todas as possibilidades de um espaço amostral determina tudo o que pode acontecer no experimento aleatório analisado.

Algumas vezes a ocorrência de um acontecimento ou de um evento pode depender ou ser alterada pela ocorrência de outro acontecimento ou evento. A próxima definição permite calcular a probabilidade de ocorrência em um experimento aleatório Σ de um evento E_i uma vez ocorrido o evento E_j.

Definição 5.2.10 (Probabilidade condicional)

A *probabilidade de ocorrência em Σ do evento $E_i(\Sigma)$, dado o evento $E_j(\Sigma)$*, denominada *probabilidade condicional de $E_i(\Sigma)$ dado $E_j(\Sigma)$* e denotada por "$p(E_i(\Sigma)|E_j(\Sigma))$", é o valor numérico definido pela seguinte equação:

$$p(E_i(\Sigma)|E_j(\Sigma)) = \frac{p(E_i(\Sigma) \cap E_j(\Sigma))}{p(E_j(\Sigma))}.$$

Decorre imediatamente da Definição acima que $p(E_i \cap E_j) = p(E_i|E_j) . p(E_j)$.

Baseados na teoria exposta nesta seção, a seguir desenvolvemos uma semântica probabilística para a Lógica Sentencial Clássica, *LSC*, tal como na Definição 2.2.16. Chamamos esta perspectiva de *semântica probabilística para LSC* (doravante, $S_{\mathbb{P}}$). Conforme mostraremos, o comportamento de $S_{\mathbb{P}}$ não é estritamente equivalente ao comportamento da semântica veritativo-funcional clássica usual (doravante $S_{\mathcal{V}}$), como desenvolvida, dentre outros, por Mendelson (1964) e Shoenfield (1967) e apresentada na Seção 3.4 deste trabalho.

3 Uma semântica probabilística para linguagens da lógica sentencial clássica

Introduzimos, inicialmente, uma associação entre as fórmulas de *L(LSC)*, como na Definição 2.2.15c, e os eventos de um experimento aleatório Σ, com um espaço amostral equiprovável $U(\Sigma) \neq \varnothing$ e finito em $\mathbb{P}$ (Definição 5.2.1). Definimos algumas noções fundamentais para nosso trabalho e mostramos uma série de resultados característicos de $S_{\mathbb{P}}$. Doravante, quando nos referirmos a uma linguagem de *LSC*, utilizaremos apenas a letra "*L*", em vez de "*L*(*LSC*). Lembramos que Form(*L*) denota o conjunto de fórmulas de *L* e que φ, ψ e γ são variáveis metalinguísticas que representam elementos de Form(*L*).

Definição 5.3.1 (Situação para uma linguagem)

Uma função *f* é uma *Σ-situação para L*, ou simplesmente uma *situação*, denotada por "$f(\Sigma)$", sse $f(\Sigma)$: Form(*L*) $\rightarrow \wp(U(\Sigma))$ tal que:

a: Se φ é atômica, então $f(\Sigma)(\varphi) = E(\Sigma)$, definido pela própria *f*.

b: Se φ é da forma $\neg\psi$, então $f(\Sigma)(\varphi) = \overline{f(\Sigma)(\psi)}$.

c: Se φ é da forma $\psi \vee \gamma$, então $f(\Sigma)(\varphi) = f(\Sigma)(\psi) \cup f(\Sigma)(\gamma)$.

Tendo em vista a Definição 5.3.1, uma situação para *L* consiste de uma atribuição de um único evento de um dado experimento aleatório a cada fórmula bem formada de *L*, de acordo com uma função *f*. O fato de uma situação ser definida a partir de uma função nos permitirá evitar ambiguidades nas próximas definições, especialmente na de consequência lógica informacional.

Dizer que *f* é uma *Σ-situação para L* corresponde a dizer que o experimento aleatório *Σ é uma f-estrutura para L*. Guardadas as devidas proporções, pode-se estabelecer uma analogia entre as noções de situação e valoração, tal como na Definição 3.4.6.

Apesar de cada fórmula de *L* ser associada a um único evento em uma dada situação $f(\Sigma)$, fórmulas distintas podem ser associadas a um mesmo evento em $f(\Sigma)$. Isso sempre ocorre, dado que o conjunto de fórmulas de *L* é infinito, diferentemente do número de eventos de um experimento aleatório, que é sempre finito.

A seguir definimos a probabilidade de uma fórmula de *L* em uma dada situação $f(\Sigma)$ para *L*.

Definição 5.3.2 (Probabilidade de fórmulas segundo uma situação)

A função *probabilidade de uma fórmula* φ segundo $f(\Sigma)$, denotada por "$P(f(\Sigma))$", é assim definida, com "p" a função probabilidade sobre eventos tal como definida em $\mathbb{P}$:

a: Se φ é atômica, $P(f(\Sigma))(\varphi) = p(f(\Sigma)(\varphi))$.

b: Se φ é da forma $\neg\psi$, $P(f(\Sigma))(\varphi) = p(\overline{f(\Sigma)(\psi)})$.

c: Se φ é da forma $\psi \vee \gamma$, $P(f(\Sigma))(\varphi) = p(f(\Sigma)(\psi) \cup f(\Sigma)(\gamma))$.

Observe-se que $P(f(\Sigma))$: $\text{Form}(L) \to [0,1] \subseteq \mathbb{Q}$, de modo que $p(f(\Sigma))(\varphi))$ é a imagem de φ segundo $P(f(\Sigma))$. O *valor probabilidade da fórmula φ segundo f(Σ)* é dado por $P(f(\Sigma))(\varphi)$, um número racional entre 0 e 1, conforme mostramos no comentário sobre o Teorema 5.2.9a. Ele é o valor probabilidade do evento de Σ correspondente a φ segundo $f(\Sigma)$.

Pode-se mostrar que $P(f(\Sigma))$ satisfaz as propriedades enunciadas no Teorema 5.2.9, interpretadas à luz de $S_{\mathbb{P}}$.

Doravante suprimiremos a expressão metalinguística “se φ é da forma ψ” e nos remeteremos diretamente à variável metalinguística “ψ”. Assim, a Definição 5.3.2c, por exemplo, seria escrita apenas como $P(f(\Sigma))(\psi \vee \gamma) = p(f(\Sigma)(\psi) \cup f(\Sigma)(\gamma))$.

A partir da Definição 2.2.6, temos que:

a: $P(\varphi \wedge \psi) = P(\neg(\neg\varphi \vee \neg\psi))$.

b: $P(\varphi \rightarrow \psi) = P(\neg\varphi \vee \psi)$.

c: $P(\varphi \leftrightarrow \psi) = P((\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi))$.

Abaixo calculamos o valor probabilidade de algumas fórmulas de *L* em duas situações dadas, baseadas nos jogos não viciados de um dado e de uma moeda.

Exemplo 5.3.3: Σ_1: $U(\Sigma_1) = \{1,2,3,4,5,6\}$; Σ_2: $U(\Sigma_2) = \{C, K\}$

φ	$f(\Sigma_1)(\varphi)$	$P(f(\Sigma_1))(\varphi)$	$f(\Sigma_2)(\varphi)$	$P(f(\Sigma_2))(\varphi)$
A_1	$\{2,4,6\} - E_1$	½	$\{C\}$	½
A_2	$\{1,3,5\} - E_2$	½	$\{K\}$	½
A_3	$\{1,2,3,5\} - E_3$	2/3	$\varnothing$	0
A_4	$\{1\} - E_4$	1/6	$\varnothing$	0
A_5	$\varnothing - E_5$	0	$\varnothing$	0
$A_1 \wedge A_2$	$\varnothing - E_1 \cap E_2$	0	$\varnothing$	0
$A_1 \wedge A_3$	$\{2\} - E_1 \cap E_3$	1/6	$\varnothing$	0
$A_2 \wedge A_3$	$\{1,3,5\} - E_2 \cap E_3$	½	$\varnothing$	0
$A_1 \vee A_3$	$\{U\} - E_1 \cup E_3$	1	$\{C\}$	½
$\neg(A_1 \vee A_3)$	$\varnothing - \overline{E_1 \cup E_3}$	0	$\{K\}$	½
$A_1 \rightarrow A_2$	$\{1,3,5\} - \bar{E}_1 \cup E_2$	½	$\{K\}$	½

Quadro 5.3.3: Exemplos da probabilidade de fórmulas em duas situações dadas.

Doravante, quando dissermos “Para toda f” queremos dizer “Para toda situação $f(\Sigma)$, dado Σ”. Também escreveremos, quando não gerar ambiguidade ou imprecisão, “$P(\varphi)$” como abreviação para “$P(f(\Sigma))(\varphi)$”, “$p(\varphi)$” como abreviação para “$p(f(\Sigma)(\varphi)$”, “$f(\varphi)$” como abreviação para “$f(\Sigma)(\varphi)$” e “f” como abreviação para “$f(\Sigma)$”.

Teorema 5.3.4: Para toda f tem-se que:

a: $P(\neg\varphi) = 1 - P(\varphi)$.

b: $f(\varphi \wedge \psi) = f(\varphi) \cap f(\psi)$.

c: $P(\varphi \wedge \psi) = p(f(\varphi) \cap f(\psi))$.

d: $P(\varphi \rightarrow \psi) = p(\overline{f(\varphi)} \cup f(\psi))$.

e: $P(\varphi \rightarrow \psi) = p(\overline{f(\varphi)}) + p(f(\varphi)) \text{ x } p(f(\psi)|f(\varphi))$.

Demonstração: Seja f uma situação para L.

a: $P(\neg\varphi) =_{\text{Def5.3.2b}}= p(\overline{f(\varphi)}) =_{\text{Teo5.2.9f}}= 1 - p(f(\varphi)) =_{\text{Def5.3.2}}= 1 - P(\varphi)$.

b: $f(\varphi \wedge \psi) =_{\text{Def2.2.6}}= f(\neg(\neg\varphi \vee \neg\,\psi)) =_{\text{Def5.3.1}}= \overline{\overline{f(\varphi)} \cup \overline{f(\psi)}} = f(\varphi) \cap f(\psi)$.

c: $P(\varphi \wedge \psi)) =_{\text{Def2.2.6}}= P(\neg(\neg\varphi \vee \neg\,\psi)) =_{\text{Def5.3.1}}= p(\overline{\overline{f(\varphi)} \cup \overline{f(\psi)}}) =_{\text{Def5.2.4}}= p(f(\varphi) \cap f(\psi))$.

d: $P(\varphi \rightarrow \psi) =_{\text{Def2.2.6}}= P(\neg\varphi \vee \psi) =_{\text{Def5.3.2}}= p(\overline{f(\varphi)} \cup f(\psi))$.

e: $P(\varphi \rightarrow \psi) =_{\text{Def2.2.6}}= P(\neg(\varphi \wedge \neg\psi)) =_{\text{Def5.3.2;Teo5.3.4c}}= p(\overline{f(\varphi) \cap \overline{f(\psi)}}) =_{\text{Teo5.2.9f}}= 1 - p(f(\varphi) \cap \overline{f(\psi)}) =_{\text{Def5.2.10}}= 1 - (p(f(\varphi)) \text{ x } p(\overline{f(\psi)}|f(\varphi))) = 1 - (p(f(\varphi)) \text{ x } 1 - p(f(\psi)|f(\varphi))) = 1 - (p(f(\varphi)) - p(f(\varphi)) \text{ x } p(f(\psi)|f(\varphi))) = 1 - p(f(\varphi)) + p(f(\varphi)) \text{ x } p(f(\psi)|f(\varphi)) = p(\overline{f(\varphi)}) + p(f(\varphi)) \text{ x } p(f(\psi)|f(\varphi))$. ■

O Teorema 5.3.4e explicita outro modo de obtermos o valor probabilidade de uma implicação material. Dele, podemos concluir que o valor probabilidade de uma fórmula implicativa material, $\varphi \rightarrow \psi$, em geral, não equivale ao valor da probabilidade condicional de $f(\psi)$, dado $f(\varphi)$, $f(\psi)|f(\varphi)$. Na maioria dos casos, $P(f)(\varphi \rightarrow \psi) \neq p(f(\psi)|f(\Sigma)(\varphi))$. Algumas exceções ocorrem quando $P(\varphi) = 1$ ou $f(\varphi) \subseteq f(\psi)$.

A seguir desenvolvemos uma análise da noção de implicação, especialmente a partir da perspectiva semântica probabilística. No transcorrer da análise, traçamos um paralelo entre os resultados oriundos desta perspectiva com a perspectiva semântica usual, $S_{\mathcal{V}}$, baseada na noção de valor de verdade.

Abordamos abaixo duas interpretações para a implicação material, uma baseada em $S_{\mathcal{V}}$ e a outra em $S_{\mathbb{P}}$. Denominamos a interpretação derivada de $S_{\mathcal{V}}$ de *implicação material veritativo-funcional* e a interpretação derivada de $S_{\mathbb{P}}$ de *implicação material probabilística*.

Tanto em $S_{\mathbb{P}}$ quanto em $S_{\mathcal{V}}$, a implicação material, $\varphi \rightarrow \psi$, possui valor *máximo* (1, no caso do valor probabilidade e V no caso do valor de verdade) em uma circunstância (situação ou valoração, conforme a Definição 3.4.6) sse φ possui valor *mínimo* (0, no caso da probabilidade e F no caso da verdade) ou ψ possui valor máximo na circunstância dada. Em sentido oposto, o valor de $\varphi \rightarrow \psi$ é mínimo sse o valor de φ é máximo e o valor de ψ é mínimo. Além disso, o valor de $\varphi \rightarrow \psi$, $\neg\varphi \vee \psi$ e $\neg(\varphi \wedge \neg\psi)$ deve ser exatamente o mesmo em qualquer circunstância.

Para $S_{\mathcal{V}}$ as características acima, definidoras da implicação material, podem ser mostradas a partir das Definições de *LSC*. Conforme já comentado, em $S_{\mathbb{P}}$ elas podem ser mostradas a partir da Definição 2.2.6 e da Definição 5.3.2, donde também resulta imediatamente que $P(\varphi \rightarrow \psi) = P(\neg\varphi \vee \psi) = P(\neg(\varphi \wedge \neg\psi))$, para toda *f*.

Dizemos que φ *implica logicamente* ψ sse, para toda *circunstância c*, o valor de φ segundo *c* é menor ou igual ao valor de ψ segundo *c*. Embora tenhamos utilizado o termo "implica" nesse caso, salientamos que ele não se refere ao conectivo da implicação material. Este conecta fórmulas, cuja conexão resulta em uma fórmula. A noção de implicação lógica é uma relação entre fórmulas e está ligada à noção de consequência lógica. No comentário seguido do Teorema 5.4.2 abaixo mostraremos uma relação entre fórmulas implicativas probabilisticamente válidas e a relação de implicação lógica.

Ainda resultam imediatamente das definições de $S_{\mathbb{P}}$ e $S_{\mathcal{V}}$ outros paralelos relativos a fórmulas: uma fórmula conjuntiva de *L* possui valor máximo sse os seus conjuntivos tiverem valor máximo; já uma fórmula disjuntiva possui valor mínimo sse os seus disjuntivos tiverem o valor mínimo.

Apesar de certas correspondências, $S_{\mathbb{P}}$ e $S_{\mathcal{V}}$ possuem diferenças semânticas substanciais. A primeira, e mais evidente, diz respeito à natureza do valor atribuído às fórmulas. Enquanto em $S_{\mathcal{V}}$ o valor é de verdade, "V" ou "F", em $S_{\mathbb{P}}$ esse valor é numérico no contínuo racional [0,1]. Mesmo sendo possível traçar um paralelo entre verdade e probabilidade, eles são conceitos distintos, com significados diferentes. Além disso, em $S_{\mathcal{V}}$ os valores possíveis de uma fórmula são apenas dois, enquanto que em $S_{\mathbb{P}}$ há uma infinidade de possíveis valores para uma fórmula. Em se tratando da implicação, essa mudança de perspectiva produz resultados específicos segundo $S_{\mathbb{P}}$, conforme mostraremos posteriormente.

Procuramos mostrar a seguir que a implicação material, também em $S_{\mathbb{P}}$, apresenta resultados paradoxais quando comparada à noção intuitiva de condicional. Para ressaltá-los, introduzimos uma concepção diferente da implicação material, denominada *implicação probabilística*.

Para tratar da concepção de implicação probabilística precisamos estender a linguagem L a fim de poder expressar esse conceito através de um novo conectivo na linguagem. Essa linguagem ampliada, denotada por "$L^{|}$", é a própria L acrescida do símbolo "|" em seu alfabeto.

A implicação probabilística é um conectivo lógico primitivo de $L^{|}$ e é denotada pelo símbolo "|". A conexão entre duas fórmulas φ e ψ de $L^{|}$ através da implicação probabilística resulta na expressão ψ|φ que, por sua vez, também é uma fórmula da linguagem estendida $L^{|}$, denominada *implicação probabilística* ou *condicional probabilística* ou *implicação probabilística de* ψ por φ. A fórmula ψ|φ deve ser lida como *φ implica probabilisticamente ψ* ou *φ é uma condição probabilística para ψ*, em que φ é denominado *antecedente* e ψ *consequente* da implicação.

Introduzimos a linguagem $L^{|}$ unicamente com o fim de comparar a implicação material, especialmente a probabilística, com a noção de implicação probabilística, visando mostrar os alcances e os limites daquela implicação. Findo este objetivo, que indicaremos no momento apropriado, voltaremos a tratar apenas da linguagem L.

As definições de $L^{|}$ são as definições de L acrescidas das cláusulas referentes ao símbolo |. Na definição de fórmula de L(*LSC*) acrescentamos a seguinte cláusula:

e: se **u** e **v** são fórmulas de $L^{|}$, então **u|v** é fórmula de $L^{|}$.

Na definição do valor probabilidade de uma fórmula em uma dada situação $f(\Sigma)$ (Definição 5.3.2), acrescentamos a seguinte cláusula, referente ao valor probabilidade de uma fórmula implicativa probabilística, denotado por "$P(f(\Sigma))(\psi|\varphi)$":

d: $P(f(\Sigma))(\psi|\varphi) = p(f(\Sigma)(\psi)|f(\Sigma)(\varphi))$.

Ao determinar o valor probabilidade das fórmulas implicativas probabilísticas, averiguamos em que medida φ interfere em ψ. Significa analisar em que medida a ocorrência de $f(\varphi)$ interfere na ocorrência de $f(\psi)$ ou, em outros termos, em que medida $f(\varphi)$ é um pressuposto para $f(\psi)$.

Os resultados expostos anteriormente sobre a probabilidade condicional entre eventos, conforme Definição 5.2.10, são igualmente aplicados à implicação probabilística, com a devida adequação conceitual e terminológica. Pode-se concluir, por exemplo, que o valor de uma fórmula implicativa probabilística com as constituintes φ e ψ nem sempre é equivalente ao valor da fórmula composta pela disjunção da negação de φ e de ψ, conforme mostraremos no final desta seção. Quando o valor do antecedente é mínimo e o do consequente é máximo em uma fórmula implicativa probabilística, o seu valor é mínimo, e não máximo, como ocorre com a implicação material, seja veritativo-funcional ou probabilística.

No seguinte exemplo definimos os valores de algumas fórmulas implicativas materiais probabilísticas e implicativas probabilísticas. Explicitamos, a partir dele, alguns resultados paradoxais (no sentido daquilo que vai além da opinião, ou seja, que não coincide com o bom senso, que contraria a intuição) referentes à implicação material probabilística. Resultados correspondentes podem ser averiguados em relação à implicação material veritativo-funcional, como mostrado na Seção 2.3, quando tratamos dos paradoxos da implicação material.

Exemplo 5.3.5: Com base no Exemplo 5.3.3, assumindo a situação $f(\Sigma_1)$, qual seja, a do jogo de dado não viciado como experimento aleatório, na qual consideramos um único lançamento do dado, ou seja, a ocorrência de apenas um acontecimento por instante de tempo, tem-se os seguintes valores probabilidade para as fórmulas abaixo (Em $f(\Sigma_1)$, $P(A_1) = ½$, $P(A_2) = ½$, $P(A_3) = 2/3$, $P(A_4) = 1/6$, $P(A_5) = 0$.):

Se φ então ψ	$P(\varphi \to \psi)$	$P(\psi\|\varphi)$	Tradução
Se A_1 então A_2	½	0	Se cai par então cai ímpar.
Se A_1 então A_3	2/3	1/3	Se cai par então cai primo.
Se A_2 então A_3	1	1	Se cai ímpar então cai primo.
Se A_5 então A_4	1	0	Se cai sete então cai um.

Quadro 5.3.5: Valor probabilidade de fórmulas condicionais.

Nas condições do exemplo acima, considerando a definição usual de número, intuímos que a ocorrência do evento cair número par exclui a possibilidade de cair número ímpar. Assim, a probabilidade de cair número par implicar na queda de ímpar deveria ser nula. Ou seja, o valor probabilidade da primeira fórmula implicativa do exemplo acima teria de ser zero, como ocorre com a condicional probabilística, mas não com a condicional material probabilística.

A segunda fórmula implicativa do exemplo acima também ilustra uma discrepância entre a implicação material probabilística e noção intuitiva de implicação. Percebemos intuitivamente que a queda de um número par restringe consideravelmente a queda de um número primo. Só haveria uma possibilidade de número par resultar em número primo. No entanto, nesse caso, diferentemente da condicional probabilística, o valor da condicional material probabilística é bastante elevado.

Na terceira fórmula do Exemplo 5.3.5 o valor probabilidade das duas condicionais coincidem, dado que, nessa circunstância, ímpar está contido em primo. Quando o antecedente da fórmula estiver contido no consequente, o valor da fórmula condicional material probabilística e da condicional probabilística é máximo. Ou seja, $f(\varphi) \subseteq f(\psi) \Rightarrow$ $P(\varphi \to \psi) = P(\psi|\varphi) = 1$. Aqui, o valor de ambas as condicionais capturam a noção intuitiva de implicação.

Por fim, a última fórmula novamente ilustra um resultado paradoxal no tocante à implicação material probabilística: uma fórmula com valor mínimo implica qualquer fórmula. No caso da implicação probabilística, o valor é mínimo, dado que segundo esta implicação, o vazio nada implica.

Outro resultado paradoxal da implicação material probabilística, quando comparada à noção intuitiva de implicação é o seguinte: considerando novamente o caso do jogo de um dado não viciado, as orações "se cai número um então cai número par" e "se cai número ímpar, então cai número par" possuem valores probabilidade distintos segundo a interpretação material probabilística. Entretanto, intuitivamente, em nenhuma das duas orações parece haver qualquer relação de implicação entre o antecedente e o consequente. Assim, o valor probabilidade de cada uma deveria ser zero, como ocorre quando essas fórmulas são interpretadas como implicações probabilísticas.

Em suma, a implicação material probabilística não captura a noção intuitiva de implicação quando esta é entendida como uma relação de causalidade clássica. Esta noção pressupõe algum tipo de conexão entre a causa (antecedente) e o efeito (consequente). No entanto, pode-se dizer que, assim como a implicação material veritativo-funcional, a implicação não expressa, e tampouco tem a intenção de fazê-lo, a noção de causalidade.

Além dos pontos levantados acima, a implicação material sofre críticas como as referentes aos conhecidos paradoxos da implicação material e à falta de relação de conteúdo entre os componentes de uma fórmula implicativa. Ao longo desta seção veremos em que medida estas críticas podem ser aplicadas à implicação material probabilística.

No próximo resultado mostramos algumas propriedades aplicadas à probabilidade de fórmulas. A partir de agora voltamos a considerar apenas a linguagem *L*.

Teorema 5.3.6

a: $P(f)(\varphi \to \psi) = 1 \Leftrightarrow f(\varphi) \subseteq f(\psi)$.

b: $P(\varphi \to \psi) = 1 \Rightarrow P(\varphi) \leq P(\psi)$.

c: $P(\varphi) = 1$ ou $P(\psi) = 1 \Rightarrow P(\varphi \vee \psi) = 1$.

d: $P(f)(\varphi \leftrightarrow \psi) = 1 \Leftrightarrow f(\varphi) = f(\psi)$.

e: $P(\varphi \leftrightarrow \psi) = 1 \Rightarrow P(\varphi) = P(\psi)$.

f: $P(\varphi \wedge \psi) = 1 \Leftrightarrow P(\varphi) = P(\psi) = 1$.

Demonstração: Seja *f* uma situação para *L*.

a: $1 =_{\mathrm{Hip}}= P(\varphi \to \psi) =_{\mathrm{Teo5.3.4d}}= p(\overline{f(\varphi)} \cup f(\psi)) =_{\mathrm{Teo5.2.9b}}= p(U) \Leftrightarrow \overline{f(\varphi)} \cup f(\psi) = U \Leftrightarrow_{\mathrm{Def5.2.4}}\Leftrightarrow f(\varphi) \subseteq f(\psi)$.

b: $P(\varphi \to \psi) = 1 \Rightarrow_{\mathrm{Teo5.3.6a}}\Rightarrow f(\varphi) \subseteq f(\psi) \Rightarrow n(f(\varphi)) \leq n(f(\psi)) \Rightarrow p(f(\varphi)) \leq p(f(\psi)) \Rightarrow P(\varphi) \leq P(\psi)$.

c: **Caso 1:** $P(\varphi) = 1$. Daí, $P(\varphi \vee \psi) =_{\mathrm{Teo5.3.4c}}= p(f(\varphi) \cup f(\psi)) = p(U \cup f(\psi)) = 1$.

Caso 2: $P(\psi) = 1$. Daí, $P(\varphi \vee \psi) =_{\mathrm{Teo5.3.4c}}= p(f(\varphi) \cup f(\psi)) = p(f(\varphi) \cup U) = 1$.

d: **(⇒):** $f(\varphi) = f(\psi) \Rightarrow_{\mathrm{Teo5.3.6a}}\Rightarrow P(\varphi \to \psi) = 1$. Mas $(f(\varphi) = f(\psi)) \Leftrightarrow (f(\psi) = f(\varphi)) \Rightarrow_{\mathrm{Teo5.3.6a}}\Rightarrow P(\psi \to \varphi) = 1$. Assim, $P(\varphi \to \psi) = P(\psi \to \varphi) = 1 \Rightarrow_{\mathrm{Teo5.3.6f}}\Rightarrow P((\varphi \to \psi) \wedge (\psi \to \varphi)) =_{\mathrm{Def2.3.1}}= P(\varphi \leftrightarrow \psi) = 1$.

(⇐): $P(\varphi \leftrightarrow \psi) =_{Def2.3.1}= P((\neg\varphi \vee \psi) \wedge (\neg\psi \vee \varphi)) =_{Hip}= 1 \Rightarrow_{Teo5.3.6f}\Rightarrow P(\neg\varphi \vee \psi) = P(\neg\psi \vee \varphi) = 1 \Rightarrow_{Def5.3.2}\Rightarrow p(\overline{f(\varphi)} \cup f(\psi)) = p(\overline{f(\psi)} \cup f(\varphi)) = 1 \Rightarrow_{Teo5.2.9b}\Rightarrow p(\overline{f(\varphi)} \cup f(\psi)) = p(\overline{f(\psi)} \cup f(\varphi)) = p(U) \Rightarrow_{Def5.2.4}\Rightarrow \overline{f(\varphi)} \cup f(\psi) = f(\psi) \cup \overline{f(\varphi)} = U \Rightarrow_{Def5.2.4}\Rightarrow f(\varphi) \subseteq f(\psi)$ e $f(\psi) \subseteq f(\varphi) \Rightarrow f(\varphi) = f(\psi)$.

e: $P(\varphi \leftrightarrow \psi) = 1 \Rightarrow_{Teo5.3.6d}\Rightarrow f(\varphi) = f(\psi) \Rightarrow p(f(\varphi)) = p(f(\psi)) \Rightarrow P(\varphi) = P(\psi)$.

f: $1 =_{Hip}= P(\varphi \wedge \psi) =_{Teo5.3.4c}= p(f(\varphi) \cap f(\psi)) =_{Teo5.2.9b}= p(U) \Leftrightarrow_{Teo5.2.8a}\Leftrightarrow n(f(\varphi) \cap f(\psi)) = n(U) \Leftrightarrow_{Def5.2.4}\Leftrightarrow f(\varphi) \cap f(\psi) = U \Leftrightarrow_{Def5.2.4}\Leftrightarrow f(\varphi) = f(\psi) = U \Leftrightarrow_{Teo5.2.8a}\Leftrightarrow p(f(\varphi)) = p(f(\psi)) = p(U) \Leftrightarrow_{Def5.3.2;\ Teo5.2.9b}\Leftrightarrow P(\varphi) = P(\psi) = 1$. ■

O Teorema 5.3.6a é próprio de $S_{\mathbb{P}}$. Ele determina que uma fórmula material implicativa possui valor probabilidade máximo em uma dada situação sse o evento associado pela situação ao antecedente da fórmula estiver contido no evento associado ao seu consequente. Este resultado não possui correspondente em $S_{\mathcal{V}}$, pois, nesta semântica, o valor de verdade de uma fórmula em uma dada valoração é definido diretamente pelo valor de verdade de suas constituintes. Não há recurso a um elemento mediador, como no caso de $S_{\mathbb{P}}$, em que o valor probabilidade de uma fórmula depende do valor probabilidade dos eventos associados às suas constituintes e não pode ser encontrado diretamente a partir do valor probabilidade delas.

O Teorema 5.3.6b pode ser enunciado nos termos de $S_{\mathcal{V}}$ e provado, tanto ele quanto a sua recíproca, para esta semântica. Se considerarmos que o valor de verdade "V" corresponde ao valor probabilidade "1" e o valor de verdade "F" corresponde ao valor probabilidade "0", em $S_{\mathcal{V}}$ pode-se mostrar o seguinte resultado: $\mathcal{V}(\varphi \rightarrow \psi) = V \Leftrightarrow \mathcal{V}(\varphi) \leq \mathcal{V}(\psi)$. No entanto, a recíproca do Teorema 5.3.6b não pode ser mostrada em $S_{\mathbb{P}}$: $P(\varphi) \leq P(\psi) \not\Rightarrow P(\varphi \rightarrow \psi) = 1$, como ilustra $f_1(\Sigma_1)$ do Exemplo 5.3.3: $P(A_1) = ½ \leq P(A_3) = 2/3$, mas $P(A_1 \rightarrow A_3) = 2/3$. Não é suficiente, para termos uma implicação material certa, do ponto de vista da probabilidade, que o seu antecedente seja menos provável que o seu consequente.

Uma vez que $P(\varphi) \leq P(\psi) \not\Rightarrow P(\varphi \rightarrow \psi) = 1$, temos também que: $P(\varphi \rightarrow \psi) \neq 1 \not\Rightarrow P(\varphi) \not\leq P(\psi)$. Quando o valor de uma fórmula material implicativa não é máximo, não podemos inferir que o valor do antecedente seja menor ou igual ao valor do consequente. Em $S_{\mathcal{V}}$ o resultado correspondente vale: $\mathcal{V}(\varphi \rightarrow \psi) = F \Rightarrow \mathcal{V}(\varphi) > \mathcal{V}(\psi)$. Este caso em particular também é um resultado de $S_{\mathbb{P}}$: $P(\varphi \rightarrow \psi) = 0 \Rightarrow P(\varphi) > P(\psi)$ dado que, nestas condições, $P(\varphi) = 1$ e $P(\psi) = 0$. O comportamento das duas perspectivas é o mesmo para este caso quando tratamos apenas de valores extremos, 0 ou 1.

Comentário semelhante ao resultado acima pode ser feito a respeito do Teorema 5.3.6c. Ele pode ser enunciado nos termos de $S_{\mathcal{V}}$ e provado, tanto ele quanto a sua recíproca, para esta semântica: $\mathcal{V}(\varphi) = V$ ou $\mathcal{V}(\psi) = V \Leftrightarrow \mathcal{V}(\varphi \vee \psi) = V$. Já no caso de $S_{\mathbb{P}}$, $P(\varphi \vee \psi) = 1 \nRightarrow P(\varphi) = 1$ ou $P(\psi) = 1$, como ilustra $f_1(\Sigma_1)$ do Exemplo 5.3.3: $P(A_1 \vee A_1) = 1$, mas $P(A_1) = ½ \leq P(A_2) = ½$. Tanto em $S_{\mathcal{V}}$ quanto em $S_{\mathbb{P}}$, de uma certeza, obtemos uma disjunção certa, tendo aquela certeza como um dos disjuntos. Mas, ao contrário de $S_{\mathcal{V}}$, em $S_{\mathbb{P}}$, dizer algo certo através de uma disjunção, não garante que pelo menos um dos disjuntos é certo. Também para este caso, quando tratamos apenas de valores extremos, 0 ou 1, temos que $P(\varphi) = 1$ ou $P(\psi) = 1 \Leftrightarrow P(\varphi \vee \psi) = 1$.

O Teorema 5.3.6d também é próprio de $S_{\mathbb{P}}$, não podendo ser enunciado nos termos de $S_{\mathcal{V}}$. Já o Teorema 5.3.6e pode ser enunciado nos termos de $S_{\mathcal{V}}$ e provado, tanto ele quanto a sua recíproca, para esta semântica: $\mathcal{V}(\varphi) = \mathcal{V}(\psi) \Leftrightarrow \mathcal{V}(\varphi \leftrightarrow \psi) = V$. Ou seja, em $S_{\mathcal{V}}$, se duas fórmulas tiverem o mesmo valor, então a sua conexão através de um bicondicional garante que o valor da fórmula resultante dessa conexão é máximo. No entanto, a recíproca do Teorema 5.3.6e não pode ser mostrada para $S_{\mathbb{P}}$: $P(\varphi) = P(\psi) \nRightarrow P(\varphi \leftrightarrow \psi) = 1$. O Exemplo 5.3.3 serve para ilustrar: $P(f(\Sigma_1))(A_1) = ½$ e $P(f(\Sigma_1))(A_2) = ½$. Nesse caso, $P(f(\Sigma_1))(A_1 \leftrightarrow A_2) = P(f(\Sigma_1))((A_1 \rightarrow A_2) \wedge (A_2 \rightarrow A_1)) = 0$. Também para este caso, quando tratamos apenas de valores extremos, 0 ou 1, temos que: $P(\varphi \leftrightarrow \psi) = 1 \Leftrightarrow P(\varphi) = P(\psi)$

Por fim, resultado correspondente ao Teorema 5.3.6f também pode ser provado em $S_{\mathcal{V}}$. Nos termos desta semântica ele é assim expressado: $\mathcal{V}(\varphi \wedge \psi) = V \Leftrightarrow \mathcal{V}(\varphi) = \mathcal{V}(\psi) = V$. Esse é o único resultado compartilhado totalmente pelas duas perspectivas semânticas que estamos investigando, independentemente dos valores dados aos constituintes da conjunção.

Em suma, o Teorema 5.3.6 explicita algumas conclusões importantes: há alguns resultados que são próprios de $S_{\mathbb{P}}$ dada a sua impossibilidade de serem enunciados nos termos de $S_{\mathcal{V}}$; há alguns resultados compartilhados pelas duas perspectivas semânticas; alguns resultados são próprios de uma ou outra perspectiva devido à natureza do valor (probabilidade ou verdade) atribuído às fórmulas; quando tratamos apenas de valores extremos (0 ou 1, correspondentes aos valores de verdade F e V), $S_{\mathbb{P}}$ e $S_{\mathcal{V}}$ compartilham relações relativas ao valor de uma fórmula e o valor de suas constituintes (esta conclusão pode ser aplicada também às fórmulas negativas, que não foram tratadas na argumentação anterior).

A seguir fazemos outra comparação entre $S_{\mathbb{P}}$ e $S_{\mathcal{V}}$, desta vez referente à relação de conteúdo entre os constituintes de uma fórmula. Em $S_{\mathcal{V}}$ há a possibilidade de ausência da relação de conteúdo entre eles, inclusive quando se trata de fórmulas verdadeiras e válidas. Mostramos que essa característica também se mantém em $S_{\mathbb{P}}$, atribuindo especial atenção à implicação material probabilística.

A combinação de espaços amostrais representa um caso no qual é permitido que as fórmulas sejam compostas por conteúdos irrelevantes ou desconexos. Suponhamos um caso em que a fórmula A_1 esteja associada ao evento "saiu cara", $\{C\}$, no lance de uma moeda e a fórmula A_2 esteja associada ao evento "saiu o número um", $\{1\}$, no lance de um dado. Pode-se juntar os dois experimentos, constituindo um novo experimento em que, por exemplo, por vezes fosse jogado um dado e por vezes fosse jogada uma moeda. Neste caso teríamos um novo experimento aleatório, com um novo espaço amostral U: $\{C, K, 1, 2, 3, 4, 5, 6\}$. Imaginemos que U seja equiprovável. Então $P(A_1 \wedge A_2) = 0$ e $P(A_1 \rightarrow A_2) = 7/8$. A fórmula conjuntiva em questão poderia ser traduzida para o Português como "saiu cara e saiu número um" e a implicativa como "se saiu cara então saiu número um". O valor probabilidade da conjunção acima parece estar ajustado ao valor intuitivo esperado. O mesmo não se pode dizer da implicação.

A combinação de espaços amostrais poderia resultar ainda em fórmulas como "se saiu cara então Érico Veríssimo é escritor", cuja probabilidade poderia variar entre zero e um. Uma sentença como esta, em que não há qualquer relação de conteúdo entre seus constituintes, pode ser considerada bastante razoável em $S_{\mathbb{P}}$.

A noção de fórmula *razoável* ou *aceitável*, *adequada*, é definida a partir do seu valor probabilidade: quanto mais próximo o valor probabilidade de uma fórmula estiver do máximo, mais razoável ela é. A seguir mostramos que nem sempre há simetria entre a razoabilidade de uma implicação probabilística e a relação de conteúdo de seus constituintes.

Teorema 5.3.7

a: $P(\varphi) \leq P(\gamma) \Rightarrow P(\varphi \to \psi) \geq P(\gamma \to \psi)$.

b: $P(\varphi) \leq P(\gamma) \Rightarrow P(\psi \to \varphi) \leq P(\psi \to \gamma)$.

Demonstração

a: $P(\varphi) \leq P(\gamma) \Rightarrow p(\varphi) \leq p(\gamma) \Rightarrow n(\varphi) \leq n(\gamma) \Rightarrow n(\overline{\varphi}) \geq n(\overline{\gamma}) \Rightarrow n(\overline{\varphi} \cup \psi) \geq n(\overline{\gamma} \cup \psi) \Rightarrow p(\overline{\varphi} \cup \psi) \geq p(\overline{\gamma} \cup \psi) \Rightarrow P(\varphi \to \psi) \geq P(\gamma \to \psi)$.

b: $P(\varphi) \leq P(\gamma) \Rightarrow p(\varphi) \leq p(\gamma) \Rightarrow n(\varphi) \leq n(\gamma) \Rightarrow n(\overline{\psi} \cup \varphi) \leq n(\overline{\psi} \cup \gamma) \Rightarrow p(\overline{\psi} \cup \varphi) \leq p(\overline{\psi} \cup \gamma) \Rightarrow P(\varphi \to \psi) \leq P(\gamma \to \psi)$.

Quando $f(\varphi) \subseteq f(\psi)$ e $f(\gamma) \subseteq f(\psi)$, temos que $P(\varphi \to \psi) = P(\gamma \to \psi)$, independentemente do valor de φ e γ. Salvo este caso, o Teorema 5.3.7a garante que, quanto menos razoável for o antecedente de uma implicação material probabilística, mais razoável é a própria implicação material probabilística. Essa razoabilidade, no entanto, é determinada unicamente em função da probabilidade do antecedente e, geralmente, não garante a relação de conteúdo entre as constituintes da implicação material, como exemplificamos a seguir.

Considerando a situação do jogo de um dado, por exemplo, pode-se dizer que a sentença “se cai número um então cai número par” é mais razoável do que a sentença “se cai número ímpar então cai par”, ainda que, em termos de relação de conteúdo, ambas pareçam iguais. Também pode ocorrer que, em uma dada situação, a sentença “se cai número par então Érico Veríssimo é escritor” seja mais razoável do que a sentença “se cai número primo então Érico Veríssimo é escritor”, embora nenhuma delas apresente qualquer relação de conteúdo entre o antecedente e o consequente. Por fim, a sentença “se a lua é de queijo então Érico Veríssimo é escritor” é totalmente razoável, ou seja, possui probabilidade um, se avaliada em uma situação em que a probabilidade da sentença “A lua é de queijo.” é zero.

O Teorema 5.3.7b mostra que o que foi dito acima sobre a razoabilidade do antecedente de uma implicação material e a razoabilidade desta implicação pode ser atribuído também ao seu consequente. Ele explicita que, quanto mais razoável for o consequente de uma implicação material probabilística, mais razoável é a própria implicação material probabilística, como ilustramos abaixo.

Considerando a situação do jogo de um dado outra vez, a sentença “se cai número par então cai número maior do que dois” é mais razoável do que “se cai número par então cai número menor ou igual a dois”, ainda que, em termos de relação de conteúdo, ambas pareçam iguais. Também pode ocorrer que, em uma dada situação, a sentença “se Érico Veríssimo é escritor então cai número primo” seja mais razoável do que a sentença “se Érico Veríssimo é escritor então cai número par”, embora nenhuma delas apresente qualquer relação de conteúdo entre o antecedente e o consequente. Por fim, “se hoje é quinta então chove ou não chove” é totalmente razoável, ou seja, possui probabilidade um, dado que o valor probabilidade de “chove ou não chove” é máximo.

Há vezes em que a razoabilidade de uma fórmula garante a relação de conteúdo entre as suas constituintes. Este é o caso da implicação material entre duas fórmulas atômicas, $A_i \rightarrow A_j$, quando $P(A_i \rightarrow A_j) = 1$, $P(A_i) \neq 0$ e $p(A_j) \neq 1$. Aqui, em primeiro lugar, conforme assegura o Teorema 5.3.6a, temos, que $f(A_i) \subseteq f(A_j)$, em segundo lugar, o evento correspondente ao antecedente é possível na situação dada e, em terceiro lugar, o evento correspondente ao consequente não é necessário nessa situação. Assim, o que está sendo dito pelo antecedente da implicação também está sendo dito pelo consequente.

Em suma, em grande número de vezes não há relação entre a razoabilidade de uma fórmula, especialmente as implicativas, e a conexão de conteúdo entre as suas constituintes. O mesmo pode ser dito para as fórmulas probabilisticamente válidas, como mostraremos na próxima seção, na qual mostramos alguns resultados gerais da Semântica Probabilística.

4 Alguns resultados da Semântica Probabilística

Apesar de possuir certas características próprias, mostraremos nesta seção que $S_{\mathbb{P}}$ é uma semântica para *LSC*. As fórmulas consideradas válidas por $S_{\mathbb{P}}$ são exatamente as mesmas consideradas válidas por $S_{\mathcal{V}}$. Para mostrar essa equivalência, partimos da definição de fórmulas válidas em $S_{\mathbb{P}}$, exposta a seguir.

Definição 5.4.1 (Inconsistência, validade e contingência de fórmulas de *L* em $S_{\mathbb{P}}$)

a: Uma fórmula φ é *probabilisticamente inconsistente* ou *probabilisticamente contraditória* sse, para toda $f(\Sigma)$, $P(f(\Sigma))(\varphi) = 0$.

b: Uma fórmula φ é *probabilisticamente válida* ou *probabilisticamente tautológica* sse, para toda $f(\Sigma)$, $P(f(\Sigma))(\varphi) = 1$.

c: Uma fórmula φ é *probabilisticamente contingente* sse não for probabilisticamente contraditória ou probabilisticamente válida.

Utilizaremos a expressão "φ é $\bot_{\mathbb{P}}$" como uma abreviação para "φ é probabilisticamente inconsistente" e a expressão "φ é $\top_{\mathbb{P}}$" como uma abreviação para "φ é probabilisticamente válida". Também utilizaremos o símbolo isolado "$\bot_{\mathbb{P}}$" como uma variável metalinguística sobre fórmulas probabilisticamente inconsistentes e o símbolo isolado "$\top_{\mathbb{P}}$" como uma variável metalinguística sobre fórmulas probabilisticamente válidas.

Utilizaremos a expressão "φ é $\bot_{\mathcal{V}}$" como uma abreviação para "φ é veritativo-funcionalmente inconsistente" e a expressão "φ é $\top_{\mathcal{V}}$" como uma abreviação para "φ é veritativo-funcionalmente válida". O símbolo isolado "$\bot_{\mathcal{V}}$" é uma variável metalinguística sobre fórmulas veritativo-funcionalmente inconsistentes e o símbolo isolado "$\top_{\mathcal{V}}$" é uma variável metalinguística sobre fórmulas veritativo-funcionalmente válidas (Tautologias).

Mostra-se facilmente, através de uma prova por indução matemática sobre o número de conectivos lógicos de uma fórmula, que as fórmulas contingentes podem apresentar qualquer valor probabilidade no contínuo racional [0,1].

Antes de averiguar alguns casos interessantes de fórmulas de *L* probabilisticamente inconsistentes ou válidas, apresentamos a seguir um teorema geral sobre as fórmulas probabilisticamente válidas.

Teorema 5.4.2

a: $P(f)(\varphi) \leq P(f)(\psi)$, para toda $f \Leftrightarrow f(\varphi) \subseteq f(\psi)$, para toda f.

b: $\varphi \rightarrow \psi$ é $\top_{\mathbb{P}} \Leftrightarrow P(\varphi) \leq P(\psi)$, para toda f.

c: $\varphi \leftrightarrow \psi$ é $\top_{\mathbb{P}} \Leftrightarrow P(\varphi) = P(\psi)$, para toda f.

Demonstração

a: **(⇒):** Seja $f_i(\Sigma_n)$ uma situação tal que $f_i(\Sigma_n)(\varphi) \nsubseteq f_i(\Sigma_n)(\psi)$, ainda que, para toda $f(\Sigma)$, $P(\varphi) \leq P(\psi)$. Se isso fosse possível, então também seria possível a existência de uma

$f_j(\Sigma_m)$ tal que $f_j(\Sigma_m)(\varphi) \nsubseteq f_j(\Sigma_m)(\psi)$. Disso resultaria que $n(f_j(\Sigma_m)(\varphi)) > n(f_j(\Sigma_m)(\psi)) \Rightarrow$ $p(f_j(\Sigma_m)(\varphi)) > p(f_j(\Sigma_m)(\psi)) \Rightarrow P(f_j(\Sigma_m))(\varphi) \nleq P(f_j(\Sigma_m))(\psi)$.

(⇐): $f(\varphi) \subseteq f(\psi)$, para toda $f \Rightarrow n(f(\varphi)) \leq n(f(\psi))$, para toda $f \Rightarrow P(f)(\varphi) \leq P(f)(\psi)$, para toda f.

b: (⇒): $\varphi \rightarrow \psi$ é $\top_{\mathbb{P}} \Rightarrow_{\text{Def5.4.1b}}\Rightarrow P(\varphi \rightarrow \psi) = 1$, para toda $f \Rightarrow_{\text{Teo5.3.6a}}\Rightarrow P(\varphi) \leq P(\psi)$, para toda f.

(⇐): $P(f)(\varphi) \leq P(f)(\psi)$, para toda $f \Rightarrow_{\text{Teo5.4.2a}}\Rightarrow f(\varphi) \subseteq f(\psi)$, para toda $f \Rightarrow_{\text{Teo5.3.6a}}\Rightarrow \varphi \rightarrow \psi$ é $\top_{\mathbb{P}}$.

c: **(⇒):** $P(\varphi \leftrightarrow \psi) = 1$, para toda $f \Rightarrow_{\text{Teo5.3.6d}}\Rightarrow P(\varphi) = P(\psi)$, para toda f.

(⇐): $P(\varphi) = P(\psi)$, para toda $f \Rightarrow_{\text{Teo5.4.2c}}\Rightarrow \varphi \rightarrow \psi$ é $\top_{\mathbb{P}}$. Temos ainda que $P(\varphi) = P(\psi)$, para toda $f \Rightarrow P(\psi) = P(\varphi)$, para toda $f \Rightarrow_{\text{Teo5.4.2c}}\Rightarrow \psi \rightarrow \varphi$ é $\top_{\mathbb{P}}$. Logo, $P(\varphi \rightarrow \psi) = P(\psi \rightarrow \varphi) = 1$, para toda $f \Rightarrow_{\text{Teo5.3.6c}}\Rightarrow P((\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)) = 1$, para toda $f \Rightarrow_{\text{Def. 2.3.1}}\Rightarrow \varphi \leftrightarrow \psi$ é $\top_{\mathbb{P}}$. ■

O primeiro resultado acima é próprio de $S_{\mathbb{P}}$ e não pode ser enunciado nos termos de $S_{\mathcal{V}}$. Já os dois últimos resultados são compartilhados pelas duas semânticas. O correspondente em $S_{\mathcal{V}}$ do Teorema 5.4.2b é: $\varphi \rightarrow \psi$ é $\top_{\mathcal{V}} \Leftrightarrow$ para toda valoração v tal que $\mathcal{V}(\varphi) = V$, tem-se que $\mathcal{V}(\psi) = V$. Esse resultado determina que $\varphi \rightarrow \psi$ é $\top_{\mathcal{V}} \Leftrightarrow \mathcal{V}(\varphi) \leq \mathcal{V}(\psi)$, para toda valoração v. Assim, em cada uma das duas perspectivas, tem-se que: *$\varphi \rightarrow \psi$ é válida sse φ implica logicamente* ψ. O resultado correspondente ao Teorema 5.4.2c é assim enunciado em $S_{\mathcal{V}}$: $\varphi \leftrightarrow \psi$ é $\top_{\mathcal{V}} \Leftrightarrow \mathcal{V}(\varphi) = \mathcal{V}(\psi)$, para toda v.

Teorema 5.4.3

a: $\varphi \wedge \neg\varphi$ é $\perp_{\mathbb{P}}$.

b: $\neg\varphi \vee \varphi$ é $\top_{\mathbb{P}}$.

c: $\varphi \rightarrow \varphi$ é $\top_{\mathbb{P}}$.

d: $P(\neg\varphi) = 1 \Leftrightarrow P(\varphi) = 0$.

Demonstração: Seja f uma situação para L.

a: $P(\varphi \wedge \neg\varphi) =_{Teo5.3.4c}= p(f(\varphi) \cap f(\neg\varphi)) =_{Def5.3.1b}= p(f(\varphi) \cap \overline{f(\varphi)}) =_{Teo5.2.9e}= 0.$

b: $P(\neg\varphi \vee \varphi) =_{Def5.3.2c;\ Def5.3.2b}= p(\overline{f(\varphi)} \cup f(\varphi)) = p(f(\varphi) \cup \overline{f(\varphi)}) =_{(Ax\mathbb{P}3)}= 1.$

c: $P(\varphi \rightarrow \varphi) =_{Teo5.3.4d}= p(\overline{f(\varphi)} \cup f(\varphi)) =_{Def5.2.4f}= p(f(\varphi)) \cup \overline{f(\varphi)}) =_{Def5.3.2}= P(\varphi \vee \neg\varphi) =_{Teo5.4.3b}= 1.$

d: $P(\neg\varphi) = 1 \Leftrightarrow_{Teo5.3.4a}\Rightarrow 1 - P(\varphi) = 1 \Leftrightarrow P(\varphi) = 0.$ ■

O teorema acima mostra que as três fórmulas correspondentes às leis aristotélicas básicas do pensamento (não-contradição, terceiro excluído e “identidade”) são válidas segundo $S_{\mathbb{P}}$. Essas três leis são, respectivamente, representadas formalmente pela negação da fórmula enunciada no Teorema 5.4.3a e pelas fórmulas enunciadas no Teorema 5.4.3b e 5.4.3c. As duas últimas são válidas segundo $S_{\mathbb{P}}$, como garantem 5.4.3b e 5.4.3c. Já a primeira é facilmente demonstrada a partir do Teorema 5.4.3d, 5.4.3a.

O Teorema 5.4.3d garante que uma fórmula é contraditória sse sua negação for válida. Assim, se ela for válida, a sua negação é contraditória. Ou seja, uma fórmula e sua negação jamais são ambas probabilisticamente válidas.

Como mostrado acima, as fórmulas correspondentes aos princípios da não-contradição e do terceiro excluído são válidas em $S_{\mathbb{P}}$. A seguir fazemos uma análise do significado dos princípios associados a essas fórmulas em $S_{\mathbb{P}}$.

De acordo com o princípio da não-contradição, algo não pode ser e não ser ao mesmo tempo. Assim, uma fórmula não pode ser verdadeira e não ser verdadeira ao mesmo tempo (segundo uma mesma circunstância), ou possuir valor máximo e não possuir valor máximo ao mesmo tempo. Entendido desse modo, $S_{\mathbb{P}}$ obedece a este princípio. Uma fórmula não pode possuir valor probabilidade n e não possuir valor probabilidade n (possuir probabilidade distinta de n) em uma mesma situação, por exemplo.

No entanto, o princípio da não-contradição possui alguns desdobramentos em $S_{\mathcal{V}}$, não satisfeitos segundo $S_{\mathbb{P}}$. Naquela semântica, pode-se dizer que, dada uma fórmula e sua negação, pelo menos uma delas é falsa (possui valor mínimo) em uma valoração; ainda, uma fórmula e sua negação jamais podem possuir o mesmo valor segundo uma mesma valoração. Esses dois desdobramentos do princípio da não-contradição não são satisfeitos em $S_{\mathbb{P}}$. Frequentemente ocorre que nem uma fórmula nem sua negação possuem valor mínimo em uma situação. Além disso, quando uma fórmula possui valor probabilidade meio, a sua negação também possui este valor, ou seja, uma fórmula e sua negação podem possuir o mesmo valor probabilidade ao mesmo tempo. Em $S_{\mathbb{P}}$ as duas variações do princípio em questão valem apenas nos casos limites, quando tratamos unicamente dos valores zero e um.

De acordo com o princípio do terceiro excluído, algo é ou não é, sem uma terceira alternativa. Assim, uma fórmula é verdadeira ou não é verdadeira (segundo uma mesma valoração), possui valor máximo ou não possui valor máximo. Entendido desse modo, $S_{\mathbb{P}}$ também obedece a este princípio, dado que uma fórmula possui valor probabilidade *n* ou não possui valor probabilidade *n* (possui valor distinto de *n*) em uma dada situação.

Como no caso da não-contradição, o terceiro excluído também possui alguns desdobramentos segundo $S_{\mathcal{V}}$ não satisfeitos em $S_{\mathbb{P}}$. Pode-se dizer, de acordo com $S_{\mathcal{V}}$, que, dada uma fórmula e sua negação, pelo menos uma delas é verdadeira (possui valor máximo) em uma valoração; ou ainda, de dois valores, uma fórmula necessariamente possui um deles. Em $S_{\mathbb{P}}$, no entanto, frequentemente ocorrem fórmulas tais que nem ela nem a sua negação possuem valor máximo em uma situação. Também é comum ocorrer que, dado um valor probabilidade *n*, nem uma dada fórmula nem sua negação possuam o valor *n*. Em $S_{\mathbb{P}}$ essas variações do princípio valem apenas nos casos limites, quando tratamos unicamente dos valores zero e um.

Embora seja possível que nem uma dada fórmula nem sua negação possuam valor máximo em uma situação, a fórmula resultante da disjunção entre eles possui o valor máximo em toda situação, como garantido pelo Teorema 5.4.3b. Em um paralelo com a linguagem natural, diríamos ser possível que nenhuma das duas sentenças “Érico Veríssimo é alto” e “Érico Veríssimo não é alto” tenha probabilidade máxima, embora “Érico Veríssimo é alto ou Érico Veríssimo não é alto” o tenha, uma vez estabelecido o conceito de “alto”. Isso ocorre devido ao fato de que, quando tomadas individualmente, as sentenças expressam parte do que pode ocorrer no mundo, enquanto a disjunção de uma sentença e sua negação expressa tudo o que pode ocorrer. Claramente, estamos pressupondo que a noção de negação também esteja sendo entendida classicamente. Assim, “Sócrates não é brasileiro” diz não apenas que ele pode ser Argentino, Grego, Chinês, mas também que pode ser homem, zebra, amarelo, casa etc..

Obviamente, $S_{\mathbb{P}}$ tampouco satisfaz o Princípio da Bivalência: há exatamente dois valores possíveis para uma fórmula de uma linguagem e cada fórmula sempre possui um e apenas um deles. Em $S_{\mathbb{P}}$ o valor probabilidade de uma fórmula pode variar no contínuo racional entre zero e um.

No Capítulo 2 classificamos os sistemas formais lógicos em clássicos e não-clássicos que, por sua vez, foram classificados em complementares e heterodoxos. A seguir propomos uma classificação semelhante aplicada às semânticas. Tomaremos como referência a semântica $S_{\mathcal{V}}$ para uma linguagem *L* de *LSC*. Assim, uma semântica será considerada heterodoxa ou complementar relativamente à semântica para *LSC* e não a partir de uma semântica para uma linguagem de um sistema formal lógico clássico em geral, como feito anteriormente.

Definição 5.4.4 (Semânticas clássicas e não-clássicas relativas a *LSC*)

a: Uma Semântica S é uma *semântica para LSC* sse o conjunto de fórmulas de *L* válidas segundo S é igual ao conjunto de fórmulas de *L* válidas segundo $S_{\mathcal{V}}$.

b: Uma Semântica S para *LSC* é uma *Semântica Clássica* sse S apresentar apenas dois valores possíveis para as fórmulas.

c: Seja S uma semântica para uma linguagem L qualquer que estende *L*. Então *S é uma Semântica Complementar* (relativa a *LSC*) sse o conjunto de fórmulas de *L* válidas segundo $S_{\mathcal{V}}$ está estritamente contido no conjunto de fórmulas válidas segundo S.

d: Seja S uma semântica para uma linguagem L qualquer que estende *L*. Então *S é uma* S*emântica Heterodoxa* (relativa a *LSC*) sse o conjunto de fórmulas de *L* válidas segundo $S_{\mathcal{V}}$ não está contido no conjunto de fórmulas de *L* válidas segundo S.

Para uma semântica S ser clássica, além de interpretar *L* e possuir o mesmo conjunto de fórmulas válidas que $S_{\mathcal{V}}$, as leis gerais aristotélicas do pensamento e seus desdobramentos são validadas por S do mesmo modo que o são por $S_{\mathcal{V}}$. No caso das semânticas heterodoxas, a definição explicita que nem toda fórmula veritativo-funcionalmente válida de *LSC* é válida segundo S. De acordo com esta classificação específica, uma semântica para um sistema formal lógico de primeira ordem clássico seria considerada uma semântica complementar.

A seguir mostramos que todo teorema de *LSC* é uma fórmula probabilisticamente válida.

Teorema 5.4.5

a: $P(\psi) = 1 \Rightarrow P(\varphi \vee \psi) = 1$.

b: $P(\varphi \vee \varphi) = 1 \Rightarrow P(\varphi) = 1$.

c: $P(\varphi \vee (\psi \vee \gamma)) = 1 \Rightarrow P((\varphi \vee \psi) \vee \gamma) = 1$.

d: $P(\gamma \vee \varphi) = P(\neg\gamma \vee \psi) = 1 \Rightarrow P(\varphi \vee \psi) = 1$.

e: $\vdash \varphi \Rightarrow \varphi$ é $\top_{\mathbb{P}}$.

Demonstração

a: $P(\varphi) =_{Def5.3.2a}= p(f(\varphi)) \leq p(f(\varphi) \cup f(\psi)) =_{Def5.3.2c}= P(\varphi \vee \psi)$. Logo, $P(\varphi \vee \varphi) = 1 \Rightarrow P(\varphi) = 1$.

b: $P(\varphi \vee \varphi) =_{Def5.3.2c}= p(f(\varphi) \cup f(\varphi)) =_{Def5.2.4f;Teo5.2.8a} = p(f(\varphi)) =_{Def5.3.2}= P(\varphi)$. Logo, $P(\varphi \vee \varphi) = 1 \Rightarrow P(\varphi) = 1$.

c: $P(\varphi \vee (\psi \vee \gamma)) =_{Def5.3.2c}= p(f(\varphi) \cup (f(\psi) \cup f(\gamma))) = p((f(\varphi) \cup f(\psi)) \cup f(\gamma)) =_{Def5.3.2c}= P((\varphi \vee \psi) \vee \gamma)$. Logo, $P(\varphi \vee (\psi \vee \gamma)) = 1 \Rightarrow P((\varphi \vee \psi) \vee \gamma) = 1$.

d: $1 =_{Hip}= P(\gamma \vee \varphi) =_{Hip}= P(\neg\gamma \vee \psi) =_{Def2.2.6}= P(\gamma \rightarrow \psi) \Rightarrow_{Teo5.3.6a}\Rightarrow f(\gamma) \subseteq f(\psi) \Rightarrow_{Hip}\Rightarrow P(\psi \vee \varphi) =_{Def5.3.2c}= p(f(\psi) \cup f(\varphi)) = p(f(\varphi) \cup f(\psi)) =_{Def5.3.2c}= P(\varphi \vee \psi) = 1$.

e: Do Teorema 5.4.3b, temos que os axiomas de *LSC* são probabilisticamente válidos. Como as regras de inferência de *LSC* preservam validade, a partir da Definição 2.2.2, tem-se que toda fórmula de uma prova é probabilisticamente válida. Como, pela Definição 2.2.3, um teorema é uma fórmula de uma prova, todo teorema é probabilisticamente válido. ■

No próximo resultado mostramos duas relações entre as duas semânticas analisadas neste trabalho. Elas serão utilizadas para mostrar a Completude Fraca e a consistência de *LSC* relativa a $S_{\mathbb{P}}$. A partir dele também mostramos que, embora $S_{\mathbb{P}}$ apresente algumas caraterísticas próprias, que a diferencia da Semântica Veritativo-Funcional, ela efetivamente é uma Semântica Clássica. Como o sistema formal lógico subjacente nesta análise é *LSC*, escreveremos "$\varphi \vdash \psi$" em vez de "$\varphi \vdash_{LSC} \psi$" como significado para "ψ é consequência lógica sintática em *LSC* de φ".

Teorema 5.4.6

a: $\mathcal{V}(\varphi) = V \Rightarrow P(f)(\varphi) = 1$, para algum f e $\mathcal{V}'(\varphi) = F \Rightarrow P(f')(\varphi) = 0$, para algum f'.

b: φ é $\top_{\mathbb{P}} \Leftrightarrow \varphi$ é $\top_{\mathcal{V}}$.

c: φ é $\top_{\mathbb{P}} \Leftrightarrow\ \vdash \varphi$.

Demonstração

a: Por indução no número n de ocorrências de conectivos primitivos em φ.

Base: n = 0. Neste caso, φ é atômica. Como toda fórmula atômica é probabilisticamente contingente, obtém-se o resultado.

Passo Indutivo: assumamos que o resultado vale para todo $j < n$.

Caso 1: φ é da forma $\neg\psi$. Pela hipótese da indução, o resultado vale para ψ, dado que possui menos conectivos lógicos do que φ.

Subcaso 1.1: Seja $\mathcal{V}(\varphi) = V \Rightarrow \mathcal{V}(\psi) = F \Rightarrow P(f)(\psi) = 0$, para algum $f \Rightarrow P(f)(\neg\psi) = 1$, para algum $f \Rightarrow P(f)(\varphi) = 1$, para algum f.

Subcaso 1.2: Seja $\mathcal{V}(\varphi) = F \Rightarrow \mathcal{V}(\psi) = V \Rightarrow P(f)(\psi) = 1$, para algum $f \Rightarrow P(f)(\neg\psi) = 0$, para algum $f \Rightarrow P(f)(\varphi) = 0$, para algum f.

Caso 2: φ é da forma $\psi \vee \gamma$. Pela hipótese da indução, o resultado vale para ψ e γ, dado que possuem menos conectivos lógicos do que φ.

Caso 2.1: $\mathcal{V}(\varphi) = V$, ou seja, $\mathcal{V}(\psi \vee \gamma) = V$.

Caso 2.1.1: $\mathcal{V}(\psi) = V \Rightarrow P(f)(\psi) = 1$, para algum $f \Rightarrow P(f)(\psi \vee \gamma) = 1$, para algum f.

Caso 2.1.2: $\mathcal{V}(\gamma) = V \Rightarrow P(f)(\gamma) = 1$, para algum $f \Rightarrow P(f)(\psi \vee \gamma) = 1$, para algum f.

Caso 2.2: $\mathcal{V}(\varphi) = F$, ou seja, $\mathcal{V}(\psi \vee \gamma) = F \Rightarrow \mathcal{V}(\psi) = F$ e $\mathcal{V}(\gamma) = F$. Daí, $P(f)(\psi) = 0$, para algum f e $P(f')(\gamma) = 0$, para algum f'. Pode-se mostrar que, neste caso, existe f'' tal que $f''(\psi) = \varnothing$ e $f''(\gamma) = \varnothing \Rightarrow P(f'')(\psi \vee \gamma) = 0$, para algum f''.

b: **(⇒):** Por absurdo, suponhamos que existe uma fórmula φ tal que φ é $\top_{\mathbb{P}}$ e φ não é $\top_{\mathcal{V}}$. Neste caso, $\mathcal{V}(\varphi) = F$, para alguma valoração $\mathcal{V} \Rightarrow_{\text{Teo5.4.6a}}\Rightarrow P(f)(\varphi) = 0$, para algum f $\Rightarrow_{\text{Def5.4.1}}\Rightarrow \varphi$ não é $\top_{\mathbb{P}}$, o que contraria a hipótese inicial.

(⇐): φ é $\top_{\mathcal{V}} \Rightarrow\ \vdash \varphi \Rightarrow_{\text{Teo5.4.5e}}\Rightarrow \varphi$ é $\top_{\mathbb{P}}$.

c: **(⇒):** φ é $\top_{\mathbb{P}} \Rightarrow_{\text{Teo5.4.6b}}\Rightarrow \varphi$ é $\top_{\mathcal{V}} \Rightarrow\ \vdash \varphi$.

(⇐): $\vdash \varphi \Rightarrow \varphi$ é $\top_{\mathcal{V}} \Rightarrow_{\text{Teo5.4.6b}}\Rightarrow \varphi$ é $\top_{\mathbb{P}}$. ■

Pode-se, por argumento semelhante, mostrar a recíproca do Teorema 5.4.6a: $P(f)(\varphi) = 1 \Rightarrow \mathcal{V}(\varphi) = V$, para alguma $\mathcal{V}$ e $P(f')(\varphi) = 0 \Rightarrow \mathcal{V}'(\varphi) = F$, para alguma $\mathcal{V}'$.

O Teorema 5.4.6b expressa que $S_{\mathbb{P}}$ *não é uma semântica heterodoxa* (relativa a *LSC*) e que $S_{\mathbb{P}}$ *não é uma semântica complementar* (relativa a *LSC*). De fato, o Teorema 5.4.6b mostra que $S_{\mathbb{P}}$ *é uma semântica para LSC*, pois o conjunto de fórmulas de *L* probabilisticamente válidas é igual ao conjunto de fórmulas de *L* veritativo-funcionalmente válidas. Embora seja uma semântica para *LSC*, $S_{\mathbb{P}}$ *não é uma semântica clássica*, de acordo com a Definição 5.4.4b. Isso ficou mostrado no comentário após o Teorema 5.4.3.

O Teorema 5.4.6c expressa o Meta-Teorema da Completude Fraco de *LSC* relativa a $S_{\mathbb{P}}$: uma fórmula φ de *L* é um teorema de *LSC* sse é uma fórmula probabilisticamente válida. A primeira parte mostra a sua *completude* de *LSC*: toda fórmula probabilisticamente válida é um teorema de *LSC*. Já a segunda parte trata da sua *correção*: todo teorema de *LSC* é uma fórmula probabilisticamente válida, ou seja, uma *fórmula válida em* $S_{\mathbb{P}}$. Esta parte poderia ter sido mostrada a partir do Teorema 5.4.3b e do Teorema 5.4.5.

Dizemos que uma situação *f é um modelo de LSC* sse, para todo axioma φ de *LSC*, $P(f)(\varphi) =1$ (Consequentemente, para todo teorema φ de *LSC* $P(f)(\varphi) =1$). Assim, temos que *toda situação é um modelo* de *LSC*. Por isso, dizemos que uma fórmula de *L* é válida segundo $S_{\mathbb{P}}$ sse é uma *fórmula probabilisticamente válida de LSC* ou *logicamente válida*.

Do Teorema 5.4.6c e do Teorema 5.4.3d explicitamos, com base em $S_{\mathbb{P}}$, que *LSC* é *consistente*: não há fórmulas tais que ela e sua negação sejam teoremas de *LSC*. Isso pode ser assim mostrado informalmente: como todo teorema de *LSC* é uma fórmula probabilisticamente válida de *LSC* e como a negação de uma fórmula probabilisticamente válida de *LSC* não é uma fórmula probabilisticamente válida de *LSC*, então a negação de um teorema de *LSC* não pode ser também um teorema de *LSC*.

Do teorema 5.4.6c resulta que os paradoxos da implicação material, conforme discutidos na Seção 2.3, também são mantidos em $S_{\mathbb{P}}$. Ou seja, as seguintes fórmulas são probabilisticamente válidas.

a: $\varphi \rightarrow (\psi \rightarrow \varphi)$ é $\top_{\mathbb{P}}$.

b: $(\varphi \rightarrow \psi) \vee (\psi \rightarrow \varphi)$ é $\top_{\mathbb{P}}$.

c: $\neg\varphi \rightarrow (\varphi \rightarrow \psi)$ é $\top_{\mathbb{P}}$.

A validade dessas três fórmulas pode ser demonstrada a partir dos termos de $S_{\mathbb{P}}$:

a: $P((\varphi \rightarrow (\psi \rightarrow \varphi))) =_{Def5.3.2}= p(\overline{f(\varphi)} \cup (\overline{f(\psi)} \cup f(\varphi))) = p((\overline{f(\varphi)} \cup f(\varphi)) \cup \overline{f(\psi)}) =_{Def5.2.4}= p(U \cup \overline{f(\psi)}) = 1$.

b: $P((\varphi \rightarrow \psi) \vee (\psi \rightarrow \varphi)) =_{Def5.3.2}= p((\overline{f(\varphi)} \cup f(\psi)) \cup (\overline{f(\psi)} \cup f(\varphi))) = p((\overline{f(\varphi)} \cup f(\varphi)) \cup (\overline{f(\psi)} \cup f(\psi))) =_{Def5.2.4}= p(U \cup U) = 1$.

c: $P(\neg\varphi \rightarrow (\varphi \rightarrow \psi)) =_{Teo5.3.4d}= p(f(\varphi) \cup \overline{f(\varphi)} \cup f(\psi)) =_{Def5.2.4}= p(U \cup f(\psi)) = 1$. ■

A seguir tratamos de um tipo especial de fórmulas implicativas materiais, aquelas em que ao menos um de seus componentes é uma fórmula probabilisticamente válida ou probabilisticamente inconsistente. Apesar de poder ser demonstrado como corolário do teorema anterior, decidimos explicitar a sua validade e demonstrá-la nos termos de $S_{\mathbb{P}}$. Na próxima seção elas serão utilizadas para mostrar algumas relações entre a quantidade de informação em fórmulas probabilisticamente válidas e a noção de consequência lógica informacional.

Teorema 5.4.7

a: $\bot_{\mathbb{P}} \rightarrow \varphi$ é $\top_{\mathbb{P}}$.

b: $\varphi \rightarrow \top_{\mathbb{P}}$ é $\top_{\mathbb{P}}$.

c: $P(\varphi \rightarrow \bot_{\mathbb{P}}) = P(\neg\varphi)$.

d: $P(\top_{\mathbb{P}} \rightarrow \varphi) = P(\varphi)$.

Demonstração: Seja *f* uma situação qualquer para *L*.

a: $P(\bot_{\mathbb{P}} \rightarrow \varphi) =_{Teo5.3.4d}= p(\overline{f(\bot_{\mathbb{P}})} \cup f(\varphi)) =_{(Ax\mathbb{P}2)}= (p(\overline{f(\bot_{\mathbb{P}})}) + p(f(\varphi))) - p(\overline{f(\bot_{\mathbb{P}})} \cap f(\varphi)) = 1 - 0 + p(f(\varphi)) - p(f(\varphi)) = 1$.

b: $P(\varphi \rightarrow \top_{\mathbb{P}}) =_{Teo5.3.4d}= p(\overline{f(\varphi)} \cup f(\top_{\mathbb{P}})) =_{(Ax\mathbb{P}2)}= (p(\overline{f(\varphi)}) + p(f(\top_{\mathbb{P}}))) - p(\overline{f(\varphi)} \cap f(\top_{\mathbb{P}})) = (1 - p(f(\varphi)) +1) - (1 - p(f(\varphi))) = 1$.

c: $P(\varphi \rightarrow \bot_{\mathbb{P}}) =_{Teo5.3.4d}= p(\overline{f(\varphi)} \cup f(\bot_{\mathbb{P}})) =_{(Ax\mathbb{P}2)}= (p(\overline{f(\varphi)}) + p(f(\bot_{\mathbb{P}}))) - p(\overline{f(\varphi)} \cap f(\bot_{\mathbb{P}})) = p(\overline{f(\varphi)}) + 0 - 0 = p(\overline{f(\varphi)})=_{Def5.3.2b}= P(\neg\varphi)$.

d: $P(\top_{\mathbb{P}} \rightarrow \varphi) =_{Teo5.3.4d}= p(\overline{f(\top_{\mathbb{P}})} \cup f(\varphi)) =_{(Ax\mathbb{P}2)}= (p(\overline{f(\top_{\mathbb{P}})}) + p(f(\varphi))) - p(\overline{f(\top_{\mathbb{P}})} \cap f(\varphi)) = 0 + p(f(\varphi)) - 0 = p(f(\varphi)) =_{Def5.3.2}= P(\varphi)$. ■

O teorema acima mostra que a semântica probabilística apresenta os mesmos resultados da semântica veritativo-funcional no tocante às fórmulas implicativas cujas partes são fórmulas probabilisticamente inconsistentes ou válidas, independentemente do valor da outra fórmula.

A seguir introduzimos a última definição de $S_{\mathbb{P}}$, que trata da equivalência probabilística entre fórmulas.

Definição 5.4.8 (Fórmulas probabilisticamente equivalentes)

Duas fórmulas, φ e ψ, são *probabilisticamente equivalentes*, denotado por "$\varphi \equiv_{\mathbb{P}} \psi$", sse, para toda situação $f(\Sigma)$, $P(f(\Sigma))(\varphi) = P(f(\Sigma))(\psi)$.

Quando não gerar ambiguidade, escreveremos "$\varphi \equiv \psi$" em vez de "$\varphi \equiv_{\mathbb{P}} \psi$". No próximo teorema mostramos alguns resultados oriundos desta última definição.

Teorema 5.4.9

a: $\varphi \equiv \psi \Leftrightarrow f(\varphi) = f(\psi)$, para toda situação f.

b: $\varphi \equiv \psi \Leftrightarrow \varphi \leftrightarrow \psi$ é $\top_{\mathbb{P}}$.

c: Se ψ surge de φ pela substituição de uma ou mais ocorrências de φ' por ψ', então $\varphi' \equiv \psi' \Rightarrow P(f)(\varphi) = P(f)(\psi)$.

Demonstração

a: **(⇒):** Sejam $\varphi \equiv \psi$ e f uma situação tal que $f(\varphi) \neq f(\psi)$. Então é possível que $n(f(\varphi)) \neq n(f(\psi)) \Rightarrow p(f(\varphi)) \neq p(f(\psi)) \Rightarrow P(\varphi) \neq P(\psi) \Rightarrow \varphi \not\equiv \psi$, contrariando a hipótese inicialmente sugerida.

(⇐): $f(\varphi) = f(\psi)$, para toda $f \Rightarrow p(f(\varphi)) = p(f(\psi))$, para toda $f \Rightarrow_{\text{Def5.3.2}}\Rightarrow P(\varphi) = P(\psi)$, para toda $f \Rightarrow_{\text{Def5.4.8}}\Rightarrow \varphi \equiv \psi$.

b: $\varphi \equiv \psi \Leftrightarrow P(\varphi) = P(\psi)$, para toda $f \Leftrightarrow_{\text{Teo5.4.2c}}\Leftrightarrow \varphi \leftrightarrow \psi$ é $\top_{\mathbb{P}}$.

c: Seja $\varphi' \equiv \psi'$. Daí, pelo Teorema 5.4.9a, $f(\varphi') = f(\psi')$. Tem-se que ψ difere de φ apenas por conter ψ' em lugares onde φ contém φ'. Como $f(\varphi') = f(\psi')$, os valores a serem calculados em ψ são exatamente os mesmos calculados em φ. Assim, $P(f)(\varphi) = P(f)(\psi)$. ■

O Teorema 5.4.9a é próprio de $S_{\mathbb{P}}$, não pode ser enunciado nos termos de $S_{\mathcal{V}}$. De acordo com ele, duas fórmulas equivalentes são associadas aos mesmos eventos em toda situação.

O Teorema 5.4.9b explicita que, para saber se duas fórmulas são probabilisticamente equivalentes, basta saber se a sua fórmula bicondicional correspondente é probabilisticamente válida. Resultado similar pode ser mostrado em $S_{\mathcal{V}}$.

O Teorema 5.4.9c determina que fórmulas com o mesmo evento associado segundo uma situação *f* são intercambiáveis em uma fórmula sem alterar o valor probabilidade em *f* da fórmula resultante.

Em $S_{\mathcal{V}}$ a substituição da ocorrência de uma subfórmula de uma fórmula por outra com o mesmo valor de verdade mantém inalterado o valor de verdade da fórmula resultante. Nessa perspectiva, fórmulas com o mesmo valor de verdade são *intercambiáveis*: a substituição de uma pela outra não altera o valor da fórmula resultante, se comparada com a original. Ou seja, pode-se mostrar em $S_{\mathcal{V}}$ o seguinte resultado, denominado *Teorema Substituição*: se ψ surge de φ pela substituição de uma ou mais ocorrências de φ' por ψ', então $\mathcal{V}(\varphi)' = \mathcal{V}(\psi') \Rightarrow \mathcal{V}(\varphi) = \mathcal{V}(\psi)$.

Em $S_{\mathbb{P}}$, no entanto, fórmulas com o mesmo valor probabilidade não são intercambiáveis. A substituição de uma fórmula pela outra com o mesmo valor probabilidade não garante que o valor probabilidade da fórmula resultante seja igual ao valor da original. Assim, não vale o *Teorema Substituição* em $S_{\mathbb{P}}$: se ψ surge de φ pela substituição de uma ou mais ocorrências de φ' por ψ', então $P(\varphi') = P(\psi') \not\Rightarrow P(\varphi) = P(\psi)$. O Exemplo 5.3.3 serve de contra-exemplo: $P(f(\Sigma_1))(A_1) = P(f(\Sigma_1))(A_2)$, mas $P(f(\Sigma_1))(A_1 \vee A_3) \neq p(f(\Sigma_1))(A_2 \vee A_3)$.

Um dos motivos pelos quais ocorre a especificidade de $S_{\mathbb{P}}$ apontada acima está no fato de que, em $S_{\mathcal{V}}$, consideramos apenas o valor V ou F das fórmulas. Já em $S_{\mathbb{P}}$ o valor de uma fórmula depende do conteúdo dos eventos associados às suas constituintes.

O próximo teorema expressa pares de fórmulas probabilisticamente equivalentes.

Teorema 5.4.10

a: $\varphi \to \psi \equiv \neg\varphi \vee \psi$.

b: $\varphi \vee \psi \equiv \neg(\neg\varphi \wedge \neg\psi)$.

c: $\varphi \wedge \psi \equiv \neg(\neg\varphi \vee \neg\psi)$.

d: $\varphi \wedge \psi \equiv \psi \wedge \varphi$.

e: $\varphi \vee \psi \equiv \psi \vee \varphi$.

f: $\neg\neg\varphi \equiv \varphi$.

g: $\varphi \wedge \varphi \equiv \varphi$.

h: $\varphi \vee \varphi \equiv \varphi$.

i: $\top_{\mathbb{P}} \wedge \varphi \equiv \varphi$.

j: $\top_{\mathbb{P}} \vee \varphi \equiv \top_{\mathbb{P}}$.

k: $\bot_{\mathbb{P}} \wedge \varphi \equiv \bot_{\mathbb{P}}$.

l: $\bot_{\mathbb{P}} \vee \varphi \equiv \varphi$

Demonstração: Seja f uma situação qualquer para L.

a: $P(\varphi \to \psi) =_{Teo5.3.4d}= p(\overline{f(\varphi)} \cup f(\psi)) =_{Def5.3.1b}= P(f(\neg\varphi) \cup f(\psi)) =_{Def5.3.2c}= P(\neg\varphi \vee \psi)$.

b: $P(\varphi \vee \psi) =_{Teo5.3.4d}= p(f(\varphi) \cup f(\psi)) =_{Def5.2.4}= p(\overline{\overline{f(\varphi)} \cap \overline{f(\psi)}}) =_{Def5.3.1b}=$
$p(\overline{(f(\neg\varphi)) \cap (f(\neg\psi))}) =_{Teo5.3.4b}= p(\overline{f(\neg\varphi \wedge \neg\psi)}) =_{Def5.3.1b}= P(\neg(\neg\varphi \wedge \neg\psi))$.

c: $\varphi \wedge \psi =_{Def2.3.1}= \neg(\neg\varphi \vee \neg\psi) \Rightarrow P(\varphi \wedge \psi) = P(\neg(\neg\varphi \vee \neg\psi))$.

d: $P(\varphi \wedge \psi) =_{Teo5.3.4c}= p(f(\varphi) \cap f(\psi)) = p(f(\psi) \cap f(\varphi)) =_{Teo5.3.4c}= P(\psi \wedge \varphi)$.

e: $P(\varphi \vee \psi) =_{Def5.3.2c}= p(f(\varphi)) \cup p(f(\psi)) = p(f(\psi)) \cup p(f(\varphi)) =_{Def5.3.2c}= P(\psi \vee \varphi)$.

f: $P(\neg\neg\varphi) =_{Def5.3.2b}= p(\overline{\overline{f(\varphi)}}) =_{Teo5.2.9f}= 1 - p(\overline{f(\varphi)}) =_{Teo5.2.9f}= 1 - (1 - p(f(\varphi))) = (1 - 1 +$
$p(f(\varphi))) = p(f(\varphi)) = P(\varphi)$.

g: $P(\varphi \wedge \varphi) =_{Teo5.3.4c} = p(f(\varphi) \cap f(\varphi)) =_{Def5.2.4g;Teo5.2.8a} = p(f(\varphi)) =_{Def5.3.2}= P(\varphi)$.

h: $P(\varphi \vee \varphi) =_{Def5.3.2c}= p(f(\varphi) \cup f(\varphi)) =_{Def5.2.4f;Teo5.2.8a} = p(f(\varphi)) =_{Def5.3.2}= P(\varphi)$.

i: $P(\top_{\mathbb{P}} \wedge \varphi) =_{Teo5.3.4c}= p(f(\top_{\mathbb{P}}) \cap f(\varphi)) = p(f(\varphi)) =_{Def5.3.2}= P(\varphi)$.

j: $P(\top_{\mathbb{P}} \vee \varphi) =_{Def5.3.2c}= p(f(\top_{\mathbb{P}}) \cup f(\varphi)) = p(f(\top_{\mathbb{P}})) =_{Def5.3.2}= P(\top_{\mathbb{P}})$.

k: $P(\bot_{\mathbb{P}} \wedge \varphi) =_{Teo5.3.4c}= p(f(\bot_{\mathbb{P}}) \cap f(\varphi)) = p(f(\bot_{\mathbb{P}})) =_{Def5.3.2}= P(\bot_{\mathbb{P}})$.

l: $P(\bot_{\mathbb{P}} \vee \varphi) =_{Def5.3.2c}= p(f(\bot_{\mathbb{P}}) \cup f(\varphi)) = p(f(\varphi)) =_{Def5.3.2}= P(\varphi)$. ■

A partir do Teorema 5.4.9a, temos que o evento, e, consequentemente, o valor probabilidade, associado em cada situação para esses pares de fórmulas é sempre o mesmo. No próximo capítulo, especialmente nos comentários sobre os Teoremas 6.2.3 e 6.2.4, analisamos as consequências deste resultado, especialmente no tocante à informação presente nas fórmulas de cada par.

O Teorema 5.4.9d e 5.4.9f explicitam que $S_{\mathbb{P}}$ não é uma semântica para a Lógica Temporal nem para a Lógica Intuicionista, conforme discutiremos no comentário em seguida ao Teorema 6.2.5.

O seguinte teorema explicita dois pares de fórmulas que não são probabilisticamente equivalentes.

Teorema 5.4.11

a: $P(\varphi \wedge \psi) \leq P(\varphi)$.

b: $P(\varphi \vee \psi) \geq P(\varphi)$.

Demonstração: Seja $f(\Sigma)$ uma situação qualquer para *L*.

a: $P(\varphi \wedge \psi) =_{Teo5.3.4c}= p(f(\varphi) \cap f(\psi)) \leq p(f(\varphi)) =_{Def5.3.2}= P(\varphi)$.

b: $P(\varphi \vee \psi) =_{Def5.3.2c}= p(f(\varphi) \cup f(\psi)) \geq p(f(\varphi)) =_{Def5.3.2a}= P(\varphi)$. ■

Em suma, neste capítulo expomos uma semântica para *LSC* a partir da noção de probabilidade das fórmulas de *L*. Propomos a definição de situação, que consiste em uma forma de associação entre fórmulas de *LSC* e eventos de um experimento aleatório, a partir do qual definimos um valor probabilidade para as fórmulas. Mostramos alguns resultados aplicados tanto a determinados tipos de fórmulas quanto à semântica probabilística em geral. Mostramos, ainda, algumas semelhanças e diferenças entre esta semântica e a veritativo-funcional. Dentre os principais resultados obtidos estão o de que o valor probabilidade de uma implicação material, em geral, não equivale ao valor da implicação probabilística. Além disso, concluímos que a implicação material, em sua versão probabilística, não captura a noção intuitiva de implicação quando esta é entendida como a relação de causalidade clássica; não captura a noção de quantidade de informação, tal como sugerida por Shannon; não supera os paradoxos da implicação material. Com relação à semântica probabilística, dentre os seus principais resultados está o de que o conjunto de fórmulas válidas segundo $S_{\mathbb{P}}$ e segundo $S_{\mathcal{V}}$ é exatamente o mesmo. Apesar disso, a semântica probabilística não é uma semântica clássica para *LSC*, dado que algumas das variações dos seus princípios básicos, como o do terceiro excluído e da não contradição, não são válidos na semântica probabilística. Concluímos ainda que $S_{\mathbb{P}}$ não captura o conteúdo informacional das sentenças.

Capítulo 6

Consequência lógica informacional

1 Apresentação

Neste capítulo sugerimos uma definição quantitativo-informacional da consequência lógica, que denominamos *consequência lógica informacional*. Introduzimos, na próxima seção, a noção de quantidade de informação presente em uma fórmula de *L*, baseados na noção de quantidade de informação desenvolvida por Shannon, exposta no Capítulo 4. Mostramos algumas relações entre o valor probabilidade e o valor informacional de uma fórmula de *L*, além de provar alguns resultados próprios da perspectiva informacional. Na terceira seção definimos o valor probabilidade de um conjunto de fórmulas para, em seguida, definir a noção de consequência lógica probabilística. Mostramos alguns resultados derivados desta definição. Por fim, na última seção, introduzimos a definição de consequência lógica informacional. Mostramos que ela se distingue das outras duas versões e que a Lógica subjacente a ela não é uma Lógica Clássica.

2 Quantidade de informação em uma fórmula

Nesta seção introduzimos a definição de quantidade de informação em uma fórmula, definida a partir do seu valor probabilidade. Mostramos, em particular, que *LSC* não captura a noção de quantidade de informação tal como definida na TMC e que tampouco captura a noção de conteúdo informacional de uma sentença.

Definição 6.2.1 (Quantidade de informação de uma fórmula segundo uma situação)
A *quantidade de informação* ou *valor informacional de uma fórmula* φ *de L segundo uma situação* $f(\Sigma)$, denotada por "$\mathrm{I}(f(\Sigma))(\varphi)$", é o valor numérico definido pela seguinte equação:

$$\mathrm{I}(f(\Sigma))(\varphi) = -\log_2 \mathrm{P}(f(\Sigma))(\varphi).$$

Quando $\mathrm{P}(f(\Sigma))(\varphi) = 0$, definimos que $-\log_2 0 = 0$, ou seja, $\mathrm{I}(f(\Sigma))(\varphi) = 0$.

Como de costume, quando não gerar ambiguidade ou imprecisão, escreveremos "I(φ)" em vez de "I(f(Σ))φ". Pode-se mostrar que I(f(Σ)): Form(L) → $\mathbb{Q}_+$.

No próximo exemplo definimos a quantidade de informação de algumas fórmulas de L, seguido de comentário a seu respeito. Para efeito didático, os exemplos sugeridos ao longo do capítulo envolvem conteúdo. No entanto, os mesmos resultados poderiam ser obtidos caso eles tivessem sido construídos em algum aparato meramente formal.

Exemplo 6.2.2: Considerando as duas situações do Exemplo 5.3.3, quais sejam, a interpretação das fórmulas de L no lance de um dado, denotada por "$f(\Sigma_1)$", e a interpretação das fórmulas de L no lance de uma moeda, denotada por "$f(\Sigma_2)$", temos as seguintes quantidades de informação para as fórmulas abaixo:

φ	$f(\Sigma_1)(\varphi)$	$I(f(\Sigma_1))\varphi$	$f(\Sigma_2)(\varphi)$	$I(f(\Sigma_2))\varphi$
A_1	{2,4,6}	1	{C}	1
A_2	{1,3,5}	1	{K}	1
A_3	{1,2,3,5}	0,58	∅	0
A_4	{1}	2,58	∅	0
$\neg A_1$	{1,3,5}	1	{K}	1
$A_1 \wedge A_3$	{2}	2,45	∅	0
$A_1 \vee A_3$	U	0	{C}	1
$A_1 \rightarrow A_2$	{1,3,5}	1	{K}	1
$A_1 \rightarrow A_3$	{1,2,3,5}	0,58	{K}	1
$A_2 \rightarrow A_3$	U	0	{C}	1

Tabela 6.2.2: Definição da quantidade de informação de fórmulas

Pode-se observar, a partir do exemplo acima, que, em cada uma das duas situações, as fórmulas A_1 e $\neg A_1$ possuem o mesmo valor informacional. Consequentemente, como no caso da probabilidade, comentado após o Teorema 5.4.3, aqui também o princípio da não-contradição não pode ser entendido classicamente. Isto por que, nos termos clássicos, uma fórmula e sua negação jamais podem possuir ambas o mesmo valor, sob a mesma circunstância, sem, com isso, gerar uma contradição. No exemplo acima, uma fórmula e sua negação possuem o mesmo valor informacional na mesma situação, sem gerar qualquer contradição.

O valor informacional da disjunção exemplificada acima é nulo, dado que a fórmula exprime tudo o que pode ocorrer no mundo, ou seja, possui probabilidade máxima. Esses casos parecem estar adequados à noção intuitiva de informação conceituada por Shannon. O mesmo não pode ser dito a respeito das fórmulas implicativas do exemplo, conforme passamos a discutir a seguir.

O valor informacional da primeira sentença implicativa material probabilística acima é a unidade pelo fato de o evento correspondente à fórmula em ambas as situações, que é o mesmo evento do seu consequente, possuir probabilidade ½. Em ambas as situações, $P(A_1 \rightarrow A_2) = P(A_2) = ½ \Leftrightarrow_{Def6.2.1}\Leftrightarrow I(A_1 \rightarrow A_2) = I(A_2) = 1$. No entanto, do ponto de vista intuitivo, o que está sendo dito nas sentenças “se cai par, então cai ímpar” e “se cai cara, então cai coroa” parece impossível de acontecer. Assim, o valor probabilidade da fórmula implicativa deveria ser zero e, consequentemente, a sua quantidade de informação também deveria ser nula. Esse paradoxo tem origem na definição de implicação material probabilística, mais especificamente no valor probabilidade definido para estas fórmulas, como já discutimos a partir do Exemplo 5.3.5.

O valor informacional da segunda sentença implicativa segundo $f(\Sigma_1)$ é menor do que o valor da fórmula anterior nesta mesma situação porque o valor informacional do consequente é menor. No Português, esta fórmula na situação dada é traduzida como: “se cai par então cai primo”. Como discutido após o Teorema 5.3.7, a diminuição do valor não se deve à relação entre as partes desta implicação material, mas ocorre simplesmente devido à alteração do valor do consequente.

A sentença do parágrafo acima ilustra que a segunda implicação material tampouco captura a noção de quantidade de informação sugerida na TMC. Intuitivamente, esta sentença diz algo bastante difícil de acontecer: a ocorrência de número primo, na hipótese de ocorrer número par. A informação nesta sentença, portanto, de acordo com a perspectiva Shannoniana, deveria ser bastante elevada, contrariamente ao que acontece no caso em questão.

Por fim, a última fórmula implicativa material, interpretada à luz de $f(\Sigma_1)$, diz que “se cai ímpar, então cai primo”. Como isso é óbvio neste caso, é correto entender que a sua quantidade de informação deve ser nula. Já de acordo com $f(\Sigma_2)$, ela diz que “se cai coroa, então cai número um" ou "se cai coroa, então a lua é de queijo". A quantidade de informação dessa sentença depende unicamente da quantidade de informação do antecedente. Este caso evidencia que a quantidade de informação de uma sentença implicativa material é independente da relação de conteúdo das suas partes.

A seguir explicitamos uma relação entre o valor probabilidade e o valor informacional de uma fórmula, como o faz Shannon (1949) com relação aos eventos de uma fonte. Propomos um exemplo que caracteriza a função informacional e mostramos que a probabilidade e o valor informacional estão numa relação inversa.

Exemplo 6.2.3: Relação entre probabilidade e quantidade de informação:

P(φ)	I(φ)
0	0
1/8	3
¼	2
1/3	1,58
½	1
2/3	0,58
¾	0,42

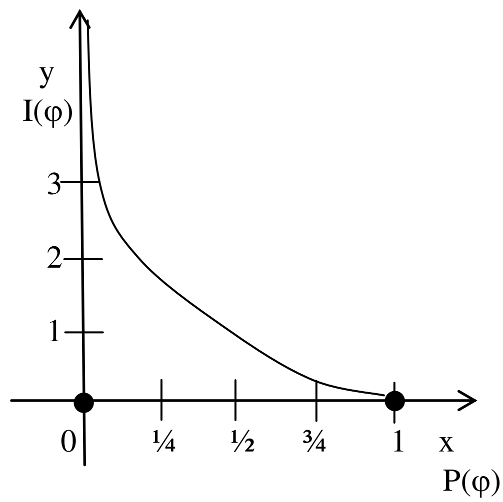


Gráfico 6.2.3: Tabela e gráfico no plano cartesiano comparando a probabilidade e quantidade de informação de uma fórmula.

O exemplo acima ilustra, tanto por meio da tabela, quanto por meio do gráfico no plano cartesiano, a relação inversa entre os valores probabilidade e informacional de uma fórmula. Quanto maior a probabilidade, menor a sua quantidade de informação e vice-versa. Nos casos limites, ou seja, probabilidade máxima e mínima, a quantidade de informação é zero. Nos demais valores probabilidade, a informação tende a zero, quando a probabilidade se aproxima do valor máximo; tende a um, quando a probabilidade se aproxima do valor mínimo.

No seguinte teorema mostramos essa relação inversa entre probabilidade e valor informacional através da comparação entre fórmulas. Também mostramos que a quantidade de informação presente em cada Teorema de *LSC* é sempre nula.

Teorema 6.2.4

a: Seja $P(\varphi) > 0$. Daí, $P(\varphi) \leq P(\psi) \Rightarrow I(\varphi) \geq I(\psi)$.

b: Seja $P(\varphi) < 1$. Daí, $I(\varphi) \geq I(\psi) \Rightarrow P(\varphi) \leq P(\psi)$.

c: $P(\varphi) = P(\psi) \Rightarrow I(\varphi) = I(\psi)$.

d: $\vdash \varphi \Rightarrow I(\varphi) = 0$, para toda situação *f*.

Demonstração: Seja *f* uma situação qualquer.

a: Sejam $P(\varphi) > 0$, $P(\varphi) \leq P(\psi)$, $P(\varphi) = p(f(\varphi)) = a/n$ e $P(\psi) = p(f(\psi)) = b/n \Rightarrow a \leq b$ e $-\log a/n \geq -\log b/n \Rightarrow I(\varphi) \geq I(\psi)$.

b: Sejam $P(\varphi) < 1$, $I(\varphi) \geq I(\psi)$, $P(\varphi) = p(f(\varphi)) = a/n$ e $P(\psi) = p(f(\psi)) = b/n \Rightarrow -\log a/n \leq -\log b/n \Rightarrow \log a \leq \log b \Rightarrow a \leq b$ e $a/n \leq b/n \Rightarrow P(\varphi) \leq P(\psi)$.

c: $P(\varphi) = P(\psi) \Rightarrow -\log_2 P(\varphi) = -\log_2 P(\psi) \Rightarrow_{Def6.2.1}\Rightarrow I(\varphi) = I(\psi)$.

d: $\vdash \varphi \Rightarrow_{Teo5.4.6c}\Rightarrow \varphi$ é $\top_{\mathbb{P}} \Rightarrow_{Def5.4.1b}\Rightarrow P(\varphi) = 1$, para toda *f* $\Rightarrow_{Def6.2.1}\Rightarrow I(\varphi) = 0$, para toda *f*. ■

Quando $P(\varphi) = 0$, sendo $0 < P(\psi) < 1$, então o Teorema 6.2.4a não vale. Neste caso, $I(\varphi) = 0$ e $I(\psi) > 0$. Ou seja, $P(\varphi) \leq P(\psi)$, mas $I(\varphi) < I(\psi)$. O mesmo pode ser dito sobre o Teorema 6.2.4b no caso em que $P(\varphi) = 1$ e $0 < P(\psi) < 1$.

Conforme explicitado pelo Teorema 6.2.4c, fórmulas com o mesmo valor probabilidade possuem o mesmo valor informacional. Embora bastante óbvio, ele expressa claramente que a perspectiva informacional em questão trata unicamente do aspecto quantitativo da informação de sentenças e não do seu conteúdo informacional, tampouco do comportamento resultante de uma informação. As fórmulas correspondentes às sentenças “caiu par” e “caiu ímpar”, no experimento aleatório que consiste no jogo de um dado não viciado, por exemplo, possuem o mesmo valor probabilidade e, consequentemente, o mesmo valor informacional, segundo este teorema. Mas, claramente, o conteúdo informacional de ambas as sentenças é radicalmente diferente. O mesmo poderíamos dizer a respeito do problema da efetividade. A ação desejada resultante a partir da informação em cada sentença pode ser completamente distinta.

A recíproca do Teorema 6.2.4c não pode ser mostrada. Juntamente com ele, elencamos abaixo alguns resultados, já provados anteriormente para $S_{\mathbb{P}}$, quando nos referíamos ao valor probabilidade de uma fórmula, que não são válidos quando aplicados à informação:

a: $I(\varphi) = I(\psi) \not\Rightarrow P(\varphi) = P(\psi)$.

b: $I(\varphi) = I(\psi)$, para toda $f \not\Rightarrow \varphi \equiv \psi$.

c: $I(\varphi) = I(\psi)$, para toda $f \not\Rightarrow \varphi \leftrightarrow \psi$ é $\top_{\mathbb{P}}$.

d: $I(\varphi) = I(\psi)$, para toda $f \not\Rightarrow \vdash \varphi \leftrightarrow \psi$.

O primeiro resultado diz que fórmulas com a mesma quantidade de informação nem sempre possuem a mesma probabilidade. Quando $P(\varphi) = 1$ e $P(\psi) = 0$, tem-se que $I(\varphi) = I(\psi)$, mas $P(\varphi) \neq P(\psi)$. Este é o caso das fórmulas probabilisticamente válidas e probabilisticamente contraditórias. Nos casos em que o valor probabilidade não é extremo, vale o resultado.

As fórmulas probabilisticamente válidas e probabilisticamente contraditórias também podem ser utilizadas para sustentar o segundo resultado, de acordo com o qual ser informacionalmente igual não significa ser probabilisticamente igual. Quando $P(\varphi) = 1$, para toda f, e $P(\psi) = 0$, para toda f, $I(\varphi) = I(\psi)$, mas $\varphi \not\equiv \psi$. Este resultado é consequência do primeiro resultado e da Definição 5.4.8.

O terceiro resultado pode ser comprovado a partir do resultado anterior e do Teorema 5.4.9b. Ele expressa que fórmulas informacionalmente iguais não garantem a validade probabilística do bicondicional constituído por eles. Isto também pode ser provado para a implicação material probabilística. O contra-exemplo construído para o segundo caso acima, em que tomamos uma fórmula probabilisticamente contraditória e outra válida, serve para ilustrar.

Do terceiro resultado e do Teorema 5.4.6c, mostramos o último resultado, segundo o qual fórmulas informacionalmente iguais não garantem que a fórmula bicondicional constituída por eles seja teorema. O mesmo pode ser dito para as fórmulas implicativas materiais probabilísticas.

Na direção inversa, todos os quatro resultados acima são válidos. O primeiro é o próprio Teorema 6.2.4c. O segundo é um caso particular do Teorema 6.2.4c: $\varphi \equiv \psi \Rightarrow I(\varphi) = I(\psi)$, para toda f. Isso significa que fórmulas probabilisticamente equivalentes são iguais informacionalmente. Como no caso do Teorema 6.2.4c, este caso também ilustra que a perspectiva informacional em questão não captura o conteúdo informacional das fórmulas. As sentenças "Chove ou não chove" e "É falso que: chove e não chove", embora sejam probabilisticamente equivalentes (mais especificamente, as suas fórmulas correspondentes em *L* são probabilisticamente equivalentes), não possuem o mesmo conteúdo informacional. Enquanto uma afirma, a outra nega. Quando as relacionamos às leis básicas aristotélicas do pensamento, isso fica ainda mais evidente: enquanto a primeira sentença diz que uma coisa é ou não é, a segunda diz que uma coisa não pode ser e não ser; na lógica clássica, a primeira pode significar que uma fórmula é verdadeira ou falsa, enquanto a segunda diria que uma fórmula não pode ser verdadeira e falsa ao mesmo tempo. A primeira consiste numa exigência e a segunda consiste numa proibição.

A "recíproca" do terceiro caso acima pode ser provada a partir da "recíproca" do segundo resultado e do Teorema 5.4.9b. Já a "recíproca" do quarto caso pode ser provada a partir da "recíproca" do terceiro e do Teorema 5.4.6c.

O Teorema 6.2.4d explicita que os Teoremas de *LSC* são informacionalmente nulos. Como a Regra de *Modus Ponens* preserva a validade probabilística, e como os axiomas de *LSC* são todos probabilisticamente válidos, os teoremas de *LSC* também são probabilisticamente válidos, ou seja, vazios informacionalmente.

A recíproca deste último resultado não vale: $I(\varphi) = 0$, para toda situação $f \not\Rightarrow\ \vdash \varphi$, dado que φ pode ser probabilisticamente contraditória e, pelo Teorema 5.4.6c, neste caso, φ não seria teorema de *LSC*. Este resultado é bastante expressivo. Em termos informacionais, não há como mostrar o Teorema da Completude de *LSC*, ou seja, o conjunto de Teoremas de *LSC* não coincide com o conjunto de fórmulas informacionalmente vazias. Essa diferença se encontra, especialmente, na incapacidade da concepção informacional em distinguir as fórmulas tautológicas das fórmulas contraditórias.

Como mostrado a partir do Gráfico 6.2.3 e do Teorema 6.2.4, a informação e a probabilidade estão em proporção inversa. No Teorema 5.3.7 mostramos que a razoabilidade (que está na mesma proporção da probabilidade) de uma implicação material, $\varphi \rightarrow \psi$, nem sempre garante a relação de conteúdo entre as suas fórmulas constituintes. Destes dois teoremas podemos concluir que:

a: Quanto mais informativo (menos razoável) for o antecedente de uma fórmula implicativa material probabilística, menos informativa (mais razoável) é a fórmula implicativa material probabilística.

b: Quanto menos informativo (mais razoável) for o consequente de uma fórmula implicativa material probabilística, menos informativa (mais razoável) é a fórmula implicativa material probabilística.

Como no caso da razoabilidade, a quantidade de informação da implicação material é determinada unicamente em função do valor informacional do antecedente e, geralmente, não garante a relação de conteúdo entre as constituintes da implicação material.

A seguir mostramos alguns pares de fórmulas cuja quantidade de informação é a mesma, independente da situação.

Teorema 6.2.5

a: $I(\varphi \rightarrow \psi) = I(\neg\varphi \vee \psi)$.

b: $I(\varphi \vee \psi) = I(\neg(\neg\varphi \wedge \neg\psi))$.

c: $I(\varphi \wedge \psi) = I(\neg(\neg\varphi \vee \neg\psi))$.

d: $I(\varphi \wedge \psi) = I(\psi \wedge \varphi)$.

e: $I(\varphi \vee \psi) = I(\psi \vee \varphi)$.

f: $I(\neg\neg\varphi) = I(\varphi)$.

g: $I(\varphi \wedge \varphi) = I(\varphi)$.

h: $I(\varphi \vee \varphi) = I(\varphi)$.

i: $I(\top_{\mathbb{P}} \wedge \varphi) = I(\varphi)$.

j: $I(\top_{\mathbb{P}} \vee \varphi) = I(\top_{\mathbb{P}})$.

k: $I(\bot_{\mathbb{P}} \wedge \varphi) = I(\bot_{\mathbb{P}})$.

l: $I(\bot_{\mathbb{P}} \vee \varphi) = I(\varphi)$.

m: $I(\top_{\mathbb{P}}) = I(\bot_{\mathbb{P}}) = 0$.

Demonstração

São demonstrados pelo Teorema 5.4.10, Definição 5.4.8 e pelo Teorema 6.2.4.c. ■

Outros casos de igualdade informacional entre fórmulas podem ser mostrados, como os da associatividade da conjunção, disjunção e biimplicação, da distributividade.

Com exceção do último item acima, no Teorema 5.4.10 mostramos que as fórmulas em cada par do Teorema 6.2.5 são equivalentes. Assim, pelo Teorema 5.4.9, a fórmula biimplicativa resultante da conexão das duas fórmulas de cada par é probabilisticamente válida. Do Teorema 5.3.6f, a fórmula implicativa resultante da conexão das duas fórmulas de cada par também é probabilisticamente válida.

O Teorema 6.2.5a é consequência imediata da própria definição de implicação material. No entanto, em termos quantitativos da informação, nem sempre seria adequado identificar o valor informacional destas duas fórmulas. Considerando a situação do jogo de um dado não viciado, a quantidade de informação na sentença “se cai primo então cai par”, por exemplo, não deveria ser igual à da sentença “não cai primo ou cai par”. De fato, a quantidade de informação da sentença implicativa, nesta situação, deveria ser bem maior do que a informação presente na sentença disjuntiva. É muito difícil, ou muito pouco provável, a ocorrência de par, na hipótese da ocorrência de primo; porém, é bem mais provável que ocorra o evento associado à sentença “não cai primo ou cair par”.

O resultado acima mostra que, ao tornar equivalentes as sentenças “$\varphi \rightarrow \psi$” e “$\neg\varphi \vee \psi$”, *LSC* não captura a noção de informação tal como definida na TMC. Ao contrário do que expressa o Teorema 6.2.5a, a redução de incerteza nessas duas fórmulas pode ser distinta. A quantidade de informação delas pode ser diferente em determinadas situações.

Além do problema da possibilidade de sentenças do tipo “$\varphi \rightarrow \psi$” e “$\neg\varphi \vee \psi$” apresentarem quantidade de informação diferente, elas também podem diferir no seu *conteúdo informacional* (grosseiramente, o que elas querem dizer). Ao proferir: “se há vida extraterrestre então eu sou trípide”, por exemplo, um falante pode não estar dizendo “não há vida extraterrestre ou eu sou um trípide”. A sentença implicativa, neste caso, significa categoricamente “não há vida extraterrestre”. Ao definir a implicação material através da disjunção e negação, ou o inverso, definir a disjunção através da implicação material e negação, *LSC* não captura nem o aspecto quantitativo nem qualitativo da informação neste tipo de fórmulas.

O segundo e terceiro casos do Teorema 6.2.5, denominados *Equivalências* ou *Leis de De Morgan*, são consequência imediata da própria Definição 2.2.6. Sobre eles também poderíamos dizer, no mínimo, que tal equivalência tampouco captura o conteúdo informacional das fórmulas. No Teorema 6.2.5c, por exemplo, a primeira fórmula é uma disjunção de duas fórmulas, enquanto a segunda fórmula é uma negação da conjunção de duas fórmulas negadas. Dizer “chove ou não chove” não significa o mesmo que “é falso que: não chove e é falso que não chove”. Discutiremos esse caso no comentário sobre o Teorema 6.2.5m abaixo.

O quarto e quinto itens do Teorema 6.2.5 tratam da comutatividade informacional da conjunção e disjunção. No caso da conjunção, esse resultado mostra que a noção de valor informacional não satisfaz a variável temporal e, portanto, assim como a Semântica Probabilística, não seria uma Semântica para a Lógica Temporal. A sentença "João abriu a porta e entrou" pode não possuir a mesma quantidade de informação do que "João entrou e abriu a porta". A probabilidade de o indivíduo entrar antes de abrir a porta pode ser muito pequena em uma dada circunstância. Já a probabilidade de ele abrir a porta e entrar, pode ser bastante grande, na mesma circunstância dada.

Considerando a ordem temporal, o conteúdo informacional dessas duas sentenças também é muito distinto. Pode-se entender que, na primeira sentença, João antes abriu a porta e depois entrou; já segunda pode significar o contrário: antes João entrou e depois abriu a porta. Outra vez, ao identificar estes pares de sentenças, concluímos que *LSC* não captura nem o aspecto quantitativo nem qualitativo da informação para estas fórmulas.

O Teorema 6.2.5f garante que a dupla negação de uma fórmula é informacionalmente equivalente à sua afirmação. Isso vai contra os princípios da Lógica Intuicionista, conforme descrita na Seção 2.3. Em termos intuitivos, esta equivalência tampouco se sustenta. Em comparação com a Língua Portuguesa, por exemplo, nem sempre negar duplamente significa afirmar. A sentença "Ele não encontrou nada", em geral não significa "Ele encontrou algo", mas de fato que "Ele nada encontrou".

Os itens g e h do Teorema em questão são conhecidos, respectivamente, como Idempotência da Conjunção e Idempotência da Disjunção. Eles expressam que a conjunção ou disjunção de uma mesma fórmula possui a mesma quantidade de informação da própria fórmula. Novamente ressaltamos o problema do conteúdo informacional, relacionado ao significado das fórmulas: a sentença "eu vou ao cinema e vou ao cinema", pode ser informacionalmente distinta da sentença "eu vou ao cinema", uma vez que, na primeira sentença, o falante reforça a sua ida ao cinema.

O Teorema 6.2.5m é um exemplo de que fórmulas informacionalmente iguais nem sempre são probabilisticamente equivalentes, como mostramos há pouco. Além disso, percebemos aqui que a única conexão entre duas fórmulas do tipo em questão através de uma implicação material resultante em uma fórmula probabilisticamente válida é $\perp_{\mathbb{P}} \rightarrow \top_{\mathbb{P}}$, como mostrado no Teorema 5.4.7.

O referido item explicita que fórmulas probabilisticamente válidas e probabilisticamente inconsistentes possuem a mesma quantidade de informação. Elas são indistinguíveis em termos informacionais. Assim, "A $\vee$ ¬A" e "A $\wedge$ ¬A" são informacionalmente iguais. Como procuramos mostrar no comentário sobre o Teorema 6.2.4c, nem sempre estas duas fórmulas possuem o mesmo conteúdo informacional. Além do que dissemos lá, podemos acrescentar que, enquanto a primeira fórmula é totalmente "vaga", diz o óbvio, que algo ocorre, a segunda é totalmente "precisa", diz o impossível, que tudo ocorre. As fórmulas probabilisticamente válidas são necessárias, associadas ao próprio espaço amostral de uma situação; já as inconsistentes são impossíveis, associadas ao conjunto vazio.

Em suma, procuramos mostrar nos comentários sobre o último teorema que, ao associar certas fórmulas, *LSC* não captura a noção de informação tal como definida na TMC, além de não capturar o conteúdo informacional das fórmulas. No próximo exemplo mostramos este mesmo ponto, agora considerando especificamente as fórmulas implicativas materiais probabilísticas, comparadas à implicação probabilística, introduzida na Seção 5.3.

Exemplo 6.2.6: Com base no Exemplo 5.3.5, assumindo a situação $f(\Sigma_1)$, qual seja, a do jogo de dado não viciado como experimento aleatório, na qual consideramos um único lançamento do dado, ou seja, a ocorrência de apenas um acontecimento em um dado instante de tempo, tem-se as seguintes quantidade de informação para as fórmulas abaixo (Lembramos que, em $f(\Sigma_1)$, $P(A_1) = ½$, $P(A_2) = ½$, $P(A_3) = 2/3$, $P(A_4) = 1/6$, $P(A_5) = 0$.):

Se φ então ψ	$P(\varphi \to \psi)$	$I(\varphi \to \psi)$	$P(\psi\|\varphi)$	$I(\psi\|\varphi)$	Tradução
Se A_1 então A_2	½	1	0	0	Se cai par, então cai ímpar.
Se A_1 então A_3	2/3	0,58	1/3	1,58	Se cai par, então cai primo.
Se A_2 então A_3	1	0	1	0	Se cai ímpar, então cai primo.
Se A_5 então A_4	1	0	0	0	Se cai sete, então cai um.

Tabela 6.2.6: Valor probabilidade e informacional de fórmulas implicativas

Em um jogo de dado, os eventos cair par e cair ímpar são mutuamente exclusivos. Quando cai par, não pode cair ímpar e vice-versa. Por isso, quando cai número par, é impossível cair número ímpar. Ou seja, “cai par” jamais implica “cai ímpar”. A primeira implicação do exemplo acima, portanto, diz algo impossível. De acordo com a noção de informação adotada na TMC, a quantidade de informação desta fórmula deveria ser zero, como ocorre com a implicação probabilística.

A segunda implicação material probabilística do exemplo acima possui menor quantidade de informação do que a primeira unicamente porque a quantidade de informação do seu consequente é menor do que a quantidade do consequente da implicação anterior. De acordo com a situação em questão, a probabilidade de cair número par implicar a queda de número primo é muito baixa, mas não impossível, como na sentença anterior. A segunda sentença é muito ousada, diz algo bastante arriscado, que elimina bastantes possibilidades. Assim, a sua probabilidade deveria ser baixa e, consequentemente, a sua quantidade de informação bastante elevada. Isso é o que ocorre com a implicação probabilística, mas não com a implicação material probabilística em questão.

Nos casos considerados neste exemplo, a ocorrência do evento ligado ao antecedente da fórmula implicativa não altera o espaço amostral subsequente. Nestes casos, ao calcular o valor de uma fórmula implicativa probabilística, $\psi|\varphi$, assumimos o evento $E_j(\Sigma)$ associado ao antecedente da fórmula como se fosse o espaço amostral de um novo experimento aleatório, chamemos $U(\Sigma')$. Já o evento $E_i(\Sigma)$ associado ao consequente da implicação é visto como se fosse transformado em um evento $E_i(\Sigma')$, constituído pelos elementos de E_i que também são elementos de $U(\Sigma')$, ou seja, $E_i(\Sigma') = E_i(\Sigma) \cap E_j(\Sigma)$. Então $p(E_i(\Sigma)|E_j(\Sigma)) = n(E_i(\Sigma'))/n(U(\Sigma')) = P(\psi|\varphi)$. Na segunda fórmula do exemplo acima, avaliamos, dentre a quantidade de números pares, quantos são primos que, neste caso, é na proporção de um para três, donde o resultado do valor probabilidade da fórmula ser 1/3. Na implicação material probabilística não há essa relação de dependência entre o antecedente e o consequente da fórmula. Esse é um dos motivos pelos quais o seu valor informacional parece não estar adequado.

A terceira implicação do Exemplo 6.2.6 afirma algo óbvio. Como todo número ímpar é primo, nesta situação, é absolutamente certo que ímpar implica primo. Assim, ambas as implicações, neste caso específico, capturam a noção de informação da TMC.

Por fim, na última sentença as duas versões da implicação possuem a mesma quantidade de informação, mas por motivos distintos. Na implicação material probabilística, esta sentença é necessária; já na versão da implicação probabilística, ela diz algo impossível. Como, em ambos os casos, não há redução de incerteza, a quantidade de informação das duas sentenças é nula. Sobre a diferença no valor probabilidade de ambas já discutimos no Exemplo 5.3.5.

Para finalizar este tópico acerca da incapacidade de *LSC* capturar a noção de informação da TMC, fazemos, a seguir, uma comparação probabilística e informacional das fórmulas implicativas materiais probabilísticas e implicativas probabilísticas em que algum de seus constituintes é probabilisticamente válido ou contraditório. Para facilitar esta comparação, apresentamos a seguinte tabela:

Exemplo 6.2.7

Ψ	$\bot_{\mathbb{P}} \to \varphi$	$\varphi \vert \bot_{\mathbb{P}}$	$\top_{\mathbb{P}} \to \varphi$	$\varphi \vert \top_{\mathbb{P}}$	$\varphi \to \bot_{\mathbb{P}}$	$\bot_{\mathbb{P}} \vert \varphi$	$\varphi \to \top_{\mathbb{P}}$	$\top_{\mathbb{P}} \vert \varphi$
$P(\psi)$	1	0	$P(\varphi)$	$P(\varphi)$	$P(\neg\varphi)$	0	1	1
$I(\psi)$	0	0	$I(\varphi)$	$I(\varphi)$	$I(\neg\varphi)$	0	0	0

Tabela 6.2.7: Comparação dos valores probabilidade e informacional de fórmulas.

Como mostrado na tabela acima, embora o significado de cada uma das implicações seja distinto, em grande parte das vezes o valor probabilidade e o valor informacional da implicação material probabilística e da implicação probabilística são o mesmo nos casos considerados.

As colunas sombreadas na tabela acima indicam as diferenças nesses valores entre as fórmulas. Na primeira região sombreada, enquanto a implicação material probabilística é necessária, a implicação probabilística é impossível. Embora, em termos probabilísticos, as duas fórmulas em questão são opostas, em termos informacionais elas possuem o mesmo valor. A primeira, por dizer algo óbvio e a segunda por dizer algo impossível. Do ponto de vista da implicação material probabilística, uma contradição implica qualquer coisa. Já do ponto de vista da implicação probabilística, uma contradição nada implica. Pode-se observar que $\varphi | \bot_{\mathbb{P}} \equiv \neg(\bot_{\mathbb{P}} \to \varphi)$.

Já na segunda região sombreada observamos uma diferença no valor probabilidade e no valor informacional das duas implicações. Como na implicação material veritativo-funcional, o valor de uma fórmula material probabilística da forma $\varphi \to \bot_{\mathbb{P}}$ depende do valor da negação de φ. Neste caso, quanto mais razoável for φ, menos razoável é a fórmula implicativa. Assim, quanto menos informativa for a fórmula φ, mais informativa será a fórmula implicativa material probabilística, $\varphi \rightarrow \perp_{\mathbb{P}}$. Já no caso da fórmula implicativa probabilística correspondente, como é impossível uma fórmula implicar probabilisticamente uma contradição, a quantidade de informação dessa fórmula é nula.

Com isto damos por encerrada a análise do valor informacional aplicado às fórmulas de *LSC*. Em especial, procuramos mostrar que *LSC* não captura a noção de quantidade de informação Shannoniana e tampouco o conteúdo informacional das suas sentenças.

A seguir introduzimos a noção de consequência lógica probabilística, voltando a tomar como referência a linguagem *L*.

3 Consequência lógica probabilística

Nesta seção definimos e analisamos a noção de consequência lógica baseada na noção de valor probabilidade definido anteriormente. Tal noção de consequência é definida a partir da comparação entre o valor probabilidade de um conjunto de fórmulas, a ser introduzido a seguir, e uma fórmula.

Definição 6.3.1 (Valor probabilidade de um conjunto de fórmulas)

Seja $\Gamma = \{\varphi_1, ..., \varphi_n\} \subseteq \mathrm{Form}(L)$. O *valor probabilidade de Γ segundo* $f(\Sigma)$, denotado por "$P^*(f(\Sigma))(\Gamma)$", é definido pela seguinte equação:

$$P^*(f(\Sigma))(\Gamma) = P(f(\Sigma))(\varphi_1 \wedge ... \wedge \varphi_n).$$

Se $\Gamma = \varnothing$, então $P^*(f(\Sigma))(\Gamma) = 1$. Quando Γ é infinito, $P^*(f(\Sigma))(\Gamma)$ é indefinido.

Observe-se que $P^*(f(\Sigma))$: $\wp(\mathrm{Form}(L)) \rightarrow [0,1] \subseteq \mathbb{Q}$, de modo que $P^*(f(\Sigma))(\Gamma))$ é a imagem de Γ segundo $P^*(f(\Sigma))$. O *valor probabilidade do conjunto de fórmulas Γ segundo $f(\Sigma)$* é dado por $P^*(f(\Sigma))(\Gamma)$, um número racional entre 0 e 1, conforme mostramos no comentário sobre o Teorema 5.2.9a. Quando não gerar ambiguidade ou imprecisão, escreveremos $P(\Gamma)$ em vez de $P^*(f(\Sigma))(\Gamma)$.

De acordo com a definição acima, o valor probabilidade de um dado conjunto de fórmulas de *L* em uma situação *f*, em última análise, é definido a partir do valor probabilidade da interseção dos eventos associados aos elementos de Γ. Ou seja, $P^*(f(\Sigma))(\Gamma)$

$=_{Def6.3.1}=$ $P(f(\Sigma))(\varphi_1 \wedge ... \wedge \varphi_n)$ $=_{Teo5.3.4c}$ $= p(f(\Sigma)(\varphi_1) \cap ... \cap f(\Sigma)(\varphi_n))$. Por conta das definições em questão, diremos que $f(\Sigma)(\Gamma) = f(\Sigma)(\varphi_1 \wedge ... \wedge \varphi_n)$.

A rigor, não há como definir o valor probabilidade de um conjunto vazio de fórmulas. A Definição 5.3.2, para a qual recorremos na Definição 6.3.1, trata do valor probabilidade apenas para fórmulas da linguagem *L*. Mas $\varnothing$ não é uma fórmula de *L*. Embora uma fórmula possa estar associada ao evento vazio, tendo seu valor probabilidade definido como zero, não faz sentido dizer que uma fórmula vazia tenha qualquer valor probabilidade.

Para não deixarmos o conjunto vazio sem uma definição do valor probabilidade, decidimos defini-lo arbitrariamente do modo acima. Podemos justificar nossa decisão do seguinte modo: dizer que $P(\Gamma) \neq 1$, neste caso, significaria a existência de alguma $\varphi_i \in \Gamma$ tal que $P(\varphi_i) \neq 1$. Como não há fórmulas em Γ com esse valor, dada a inexistência de fórmulas em Γ, tem-se que $P(\Gamma) = 1$.

A próxima definição delimita algumas características de um conjunto de fórmulas.

Definição 6.3.2 (Propriedades sobre conjuntos de fórmulas)

a: Uma *situação f(Σ) é um modelo de um conjunto* Γ *de fórmulas* sse, para toda $\varphi \in \Gamma$, $P(f(\Sigma))(\varphi) = 1$.

b: Uma *fórmula* φ *é probabilisticamente certa na situação f(Σ)* sse $\Gamma = \{\varphi\}$ e $f(\Sigma)$ é modelo de Γ.

c: Um *conjunto* Γ *de fórmulas é satisfatível* sse existir uma situação $f(\Sigma)$ tal que $f(\Sigma)$ é modelo de Γ.

d: Um *conjunto* Γ *de fórmulas é contraditório* sse existir uma fórmula φ tal que $\varphi \in \Gamma$ e $\neg\varphi \in \Gamma$.

São consequências imediatas da Definição 6.3.2 que:

- Uma fórmula φ é probabilisticamente certa em $f(\Sigma) \Leftrightarrow P(f(\Sigma))(\varphi) = 1$.
- Uma situação $f(\Sigma)$ é modelo de $\Gamma \Leftrightarrow$ para toda $\varphi \in \Gamma$, φ é probabilisticamente certa em $f(\Sigma)$.
- Uma fórmula φ é probabilisticamente certa em um modelo $f(\Sigma)$ de $\Gamma \Leftrightarrow P(f(\Sigma))(\varphi) = 1$.

Isto quer dizer que: as fórmulas probabilisticamente certas em uma situação possuem valor máximo nela, como era de se esperar; as fórmulas de um conjunto Γ são probabilisticamente certas em cada modelo de Γ, ou seja, f também é modelo de $\{\varphi\}$, para toda $\varphi \in \Gamma$; qualquer fórmula probabilisticamente certa em um modelo de Γ possui probabilidade certa nessa situação.

Os três primeiros casos da Definição 6.3.2 são referentes à parte semântica de um sistema formal lógico. Já a quarta é uma definição que diz respeito ao seu estudo sintático. No próximo teorema fazemos algumas associações entre tais concepções, relacionando os conceitos da semântica aos conceitos da sintaxe.

Teorema 6.3.3

a: $f(\Sigma)$ é modelo de $\Gamma \Leftrightarrow P(f(\Sigma))(\Gamma) = 1$.

b: Γ é contraditório $\Rightarrow$ Γ é insatisfatível.

c: Γ é contraditório $\Rightarrow P(f(\Sigma))(\Gamma) = 0$, para toda $f(\Sigma)$.

Demonstração: Seja f uma situação qualquer.

a: f é modelo de $\Gamma = \{\varphi_1, ..., \varphi_n\} \Leftrightarrow_{\text{Def6.3.2a}}\Leftrightarrow P(f)(\varphi_1) = ... = P(f)(\varphi_n) = 1 \Leftrightarrow_{\text{Teo5.3.6f}}\Leftrightarrow P(f)(\varphi_1 \wedge ... \wedge \varphi_n) = 1 \Leftrightarrow_{\text{Def6.3.1}}\Leftrightarrow P(f)(\Gamma) = 1$.

b: Γ é contraditório $\Rightarrow_{\text{Def6.3.2d}}\Rightarrow$ existe φ tal que $\varphi, \neg\varphi \in \Gamma = \{\varphi, \neg\varphi, ..., \varphi_n\}$, para $n \geq 2$. Mas $P(\neg\varphi) = 1 \Leftrightarrow_{\text{Teo5.4.3d}}\Leftrightarrow P(\varphi) = 0$. Assim, $P(\neg\varphi) \neq 1$ ou $P(\varphi) \neq 1$, para toda f $\Rightarrow_{\text{Teo5.3.6f}}\Rightarrow P(f)(\Gamma) \neq 1$, para toda f $\Rightarrow_{\text{Def6.3.2a}}\Rightarrow$ f não é modelo de Γ, para toda f $\Rightarrow_{\text{Def6.3.2c}}\Rightarrow$ Γ é insatisfatível.

c: Γ é contraditório $\Rightarrow_{\text{Def6.3.2d}}\Rightarrow$ existe φ tal que $\varphi, \neg\varphi \in \Gamma = \{\varphi, \neg\varphi, ..., \varphi_n\}$, para $n \geq 2$ $\Rightarrow_{\text{Def6.3.1}}\Rightarrow P(\Gamma) = P(\varphi \wedge \neg\varphi \wedge ... \wedge \varphi_n) \Rightarrow_{\text{Teo5.4.10k}}\Rightarrow P(\Gamma) = 0$, para toda f. ■

O Teorema 6.3.3a diz que o valor probabilidade de um conjunto de fórmulas é máximo em toda situação que é seu modelo. O Teorema 6.3.3b estabelece que um conjunto de fórmulas contraditório não possui modelos. Em nenhuma situação todas as fórmulas de um conjunto possuem probabilidade máxima quando ele for contraditório. Assim, se um conjunto de fórmulas possuir um modelo, para qualquer fórmula, ou ela ou a sua negação não pertencem a ele. Ou seja, $f(\Sigma)$ é modelo de $\Gamma \Rightarrow$ para toda φ, $\varphi \notin \Gamma$ ou $\neg\varphi \notin \Gamma$. Já o terceiro item do teorema acima explicita que os conjuntos contraditórios possuem probabilidade nula.

A recíproca de cada um dos dois últimos resultados do teorema acima não pode ser mostrada em $S_{\mathbb{P}}$. Ou seja:

a: Γ é insatisfatível $\not\Rightarrow$ Γ é contraditório.

b: $P(\Gamma) = 0$, para toda $f \not\Rightarrow \Gamma$ é contraditório.

Para mostrar o primeiro caso, apresentamos o seguinte conjunto: $\Gamma = \{\varphi, \varphi \rightarrow \psi, \neg\psi\}$. Este conjunto não é contraditório, embora seja insatisfatível, como percebemos através da seguinte argumentação: suponha que f seja um modelo de $\Gamma \Rightarrow_{Def5.4.2a} \Rightarrow P(\varphi) = P(\neg\psi) = 1 \Rightarrow_{Teo5.3.4.a} \Rightarrow P(\neg\varphi) = P(\neg\neg\psi) =_{Teo5.4.10f} = P(\psi) =_{Teo5.4.3d} = 0 \Rightarrow_{Def5.3.2} \Rightarrow P(\varphi \rightarrow \psi) = 0$. Mas $P(\varphi \rightarrow \psi) = 1 \Rightarrow_{Teo5.3.4d;Def5.3.2c} = P(\varphi) \neq 1$ ou $P(\psi) \neq 0 \Leftrightarrow_{Teo5.4.3d} \Leftrightarrow P(\neg\psi) \neq 1$. Logo, f não é modelo de Γ.

O seguinte exemplo serve para mostrar o segundo item acima: seja $\Gamma = \{(\varphi \wedge \neg\varphi)\} \Rightarrow_{Def6.3.1} \Rightarrow P(\Gamma) = 0$, para toda f. No entanto, o conjunto Γ não é contraditório, embora contenha, como único elemento, uma fórmula probabilisticamente contraditória. Isto ilustra a diferença entre fórmula probabilisticamente contraditória e conjunto contraditório de fórmulas.

A Definição 6.3.3 não pode ser confundida com as noções de *Teoria Insatisfatível* (que não tem modelos) e *Teoria Inconsistente* (aquela na qual toda fórmula da sua linguagem é um teorema da teoria), como definidas usualmente para a Lógica Clássica. De acordo com estas definições, pode-se mostrar que uma teoria é consistente sse for satisfatível. Nas Considerações Finais deste trabalho mostramos que, de acordo com a definição de consequência lógica informacional, uma teoria pode ser inconsistente (seu conjunto de Teoremas é contraditório) sem com isso ser trivial (que qualquer fórmula da linguagem pode ser considerada um teorema da teoria).

Dados estes conceitos introdutórios, apresentamos a seguir a definição de consequência lógica probabilística.

Definição 6.3.4 (Consequência lógica probabilística)

Uma fórmula *φ é consequência lógica probabilística* de um conjunto Γ de fórmulas, denotado por "$\Gamma \vDash_{\mathbb{P}} \varphi$", sse, para toda situação $f(\Sigma)$, $P(f(\Sigma))(\Gamma) \leq P(f(\Sigma))(\varphi)$.

Quando $\Gamma = \varnothing$, em vez de $\Gamma \vDash_{\mathbb{P}} \varphi$ escrevemos simplesmente $\vDash_{\mathbb{P}} \varphi$. A expressão "$\Gamma \nvDash_{\mathbb{P}} \varphi$" denota que φ não é consequência lógica probabilística de Γ. As fórmulas de Γ são chamadas de *premissas* e φ é denominada de *conclusão*. Denotaremos o conjunto de consequências probabilísticas de Γ por $\vDash_{\mathbb{P}}(\Gamma)$, ou seja: $\vDash_{\mathbb{P}}(\Gamma) = \{\varphi \mid \Gamma \vDash_{\mathbb{P}} \varphi\}$. Diremos que uma fórmula é *satisfeita* sse o evento associado a ela pela situação é o evento efetivamente ocorrido no experimento aleatório correspondente.

A Definição 6.3.4 expressa que quando uma fórmula φ é consequência lógica probabilística de um conjunto Γ de fórmulas, é impossível a probabilidade de Γ ser maior do que a probabilidade de φ. Esta impossibilidade pode ser um indicativo de que a consequência lógica probabilística satisfaz a característica da Necessidade da consequência lógica, tal como descrito na Seção 1.2. Nas Considerações Finais deste trabalho voltaremos a este assunto.

A seguir apresentamos alguns pares de esquemas de fórmulas de *L*(*LSC*) tal que uma é consequência lógica probabilística da outra. Este resultado poderia ser provado como um corolário dos próximos teoremas. Decidimos inserí-lo com o fim de explicitar algumas destas consequências que, em discussões futuras, serão utilizadas para analisar a adequação intuitiva tanto desta definição de consequência lógica quanto da definição informacional de consequência lógica. Como o sistema formal lógico subjacente nesta análise é *LSC*, escreveremos "$\varphi \vDash_{\mathbb{P}} \psi$" em vez de "$\varphi \vDash_{\mathbb{P}LSC} \psi$" significando que "$\psi$ é consequência lógica probabilística em *LSC* de φ".

Teorema 6.3.5

a: $\varphi \rightarrow \psi \vDash_{\mathbb{P}} \neg\varphi \vee \psi$.

b: $\varphi \vee \psi \vDash_{\mathbb{P}} \neg(\neg\varphi \wedge \neg\psi)$.

c: $\varphi \wedge \psi \vDash_{\mathbb{P}} \neg(\neg\varphi \vee \neg\psi)$.

d: $\varphi \wedge \psi \vDash_{\mathbb{P}} \psi \wedge \varphi$.

e: $\varphi \vee \psi \vDash_{\mathbb{P}} \psi \vee \varphi$.

f: $\neg\neg\varphi \vDash_{\mathbb{P}} \varphi$.

g: $\varphi \wedge \varphi \vDash_{\mathbb{P}} \varphi$.

h: $\varphi \vee \varphi \vDash_{\mathbb{P}} \varphi$.

i: $\top_{\mathbb{P}} \wedge \varphi \vDash_{\mathbb{P}} \varphi$.

j: $\top_{\mathbb{P}} \vee \varphi \vDash_{\mathbb{P}} \top_{\mathbb{P}}$.

k: $\bot_{\mathbb{P}} \wedge \varphi \vDash_{\mathbb{P}} \bot_{\mathbb{P}}$.

l: $\bot_{\mathbb{P}} \vee \varphi \vDash_{\mathbb{P}} \varphi$.

m: $\top_{\mathbb{P}} \vDash_{\mathbb{P}} \bot_{\mathbb{P}}$.

n: $\varphi \vDash_{\mathbb{P}} \varphi$.

Demonstração

São demonstrados pelo Teorema 5.4.10, Definição 5.4.8 e pela Definição 6.3.4. ■

A recíproca de cada um dos itens do teorema acima também pode ser mostrada, por meio da mesma base demonstrativa.

Teorema 6.3.6: Seja $\Gamma = \{\varphi_1, ..., \varphi_n\}$ o conjunto de premissas de cada item abaixo e seja φ_m a sua conclusão, para n, $m \in \mathbb{N}$. Então:

a: $\varphi \wedge \psi \vDash_{\mathbb{P}} \varphi$.

b: $\varphi \rightarrow \psi, \varphi \vDash_{\mathbb{P}} \psi$.

c: $\varphi \rightarrow \psi, \neg\psi \vDash_{\mathbb{P}} \neg\varphi$.

d: $\varphi \rightarrow \psi, \psi \rightarrow \gamma \vDash_{\mathbb{P}} \varphi \rightarrow \gamma$.

e: $\varphi \vDash_{\mathbb{P}} \psi \rightarrow \varphi$.

f: $\psi \vDash_{\mathbb{P}} \varphi \vee \psi$.

g: $\varphi \vee \varphi \vDash_{\mathbb{P}} \varphi$.

h: $\varphi \vee (\psi \vee \gamma) \vDash_{\mathbb{P}} (\varphi \vee \psi) \vee \gamma$.

i: $\gamma \vee \varphi, \neg\gamma \vee \psi \vDash_{\mathbb{P}} \varphi \vee \psi$.

Demonstração: seja *f* uma situação qualquer.

Pode-se mostrar, para cada item, que $f(\varphi_1 \wedge ... \wedge \varphi_n) \subseteq f(\varphi_m) \Rightarrow_{Teo5.4.2a}\Rightarrow P(\varphi_1 \wedge ... \wedge \varphi_n) =_{Def6.3.1}= P(\Gamma) \leq P(\varphi_m) \Rightarrow_{Def6.3.4}\Rightarrow \Gamma \vDash_{\mathbb{P}} \varphi_m$. ■

Os itens do teorema acima são correspondentes a algumas das regras de dedução comumente adotados em certos sistemas formais lógicos. O primeiro item é a eliminação da conjunção; o segundo, *Modus Ponens*; o terceiro é *Modus Tollens*; o quarto é a transitividade; o quinto é a introdução da implicação. Os quatro últimos itens são correspondentes às regras de inferência de *LSC*.

É facilmente mostrado que, para cada caso do teorema acima, $P(\varphi_1) = ... = P(\varphi_n) = 1 \Rightarrow P(\varphi_m) = 1$. Ou seja, as regras enunciadas preservam a validade probabilística.

No próximo teorema mostramos algumas associações gerais da consequência lógica probabilística com outras noções. Ela é comparada a noções pertencentes à Semântica Probabilística exposta na seção anterior e relacionada à consequência lógica sintática, e em particular à noção de Teorema de *LSC*, donde mostramos uma relação dela com a consequência lógica funcional-veritativa.

Como o sistema formal lógico subjacente nesta análise é *LSC*, escreveremos "$\varphi \vDash_{\mathcal{V}} \psi$" em vez de "$\varphi \vDash_{\mathcal{V}LSC} \psi$" para significar que ψ é consequência lógica veritativo-funcional em *LSC* de φ.

Teorema 6.3.7

a: $\vDash_{\mathbb{P}} \varphi \Leftrightarrow \varphi$ é $\top_{\mathbb{P}}$.

b: $\Gamma \vDash_{\mathbb{P}} \varphi \Leftrightarrow P(\Gamma) = 1 \Rightarrow P(\varphi) = 1$, para toda f.

c: $\Gamma \vDash_{\mathbb{P}} \varphi \Leftrightarrow f(\Gamma) \subseteq f(\varphi)$, para toda f.

d: $\Gamma \vDash_{\mathbb{P}} \varphi \Leftrightarrow \Gamma \vDash_{\mathcal{V}} \varphi$.

e: $\Gamma \vDash_{\mathbb{P}} \varphi \Leftrightarrow \Gamma \vdash \varphi$.

Demonstração

a: $\vDash_{\mathbb{P}} \varphi \Leftrightarrow_{\text{Def6.3.4}}\Leftrightarrow P(\varnothing) \leq P(\varphi)$, para toda f $\Leftrightarrow_{\text{Def6.2.7}}\Leftrightarrow P(\varphi) = 1$, para toda f $\Leftrightarrow_{\text{Def5.4.1b}}\Leftrightarrow \varphi$ é $\top_{\mathbb{P}}$.

b: ($\Rightarrow$): $\Gamma \vDash_{\mathbb{P}} \varphi \Rightarrow_{\text{Def6.3.4}}\Rightarrow P(\Gamma) \leq P(\varphi)$, para toda $f \Rightarrow P(\Gamma) = 1 \Rightarrow P(\varphi) = 1$, para toda f.

($\Leftarrow$): suponhamos que, para toda f, $P(\Gamma) = 1 \Rightarrow P(\varphi) = 1$ e que $\Gamma \nvDash_{\mathbb{P}} \varphi \Rightarrow_{\text{Def6.3.4}}\Rightarrow P(\Gamma) > P(\varphi)$, para alguma f'. Sejam $P(f)(\Gamma) = 1$ e $f(\varphi) = f'(\varphi) \Rightarrow P(f)(\Gamma) = 1$ e $P(f)(\varphi) < 1$.

c: $\Gamma \vDash_{\mathbb{P}} \varphi \Leftrightarrow_{\text{Def6.3.4}}\Leftrightarrow P(\Gamma) \leq P(\varphi)$, para toda $f \Leftrightarrow_{\text{Teo5.4.2a}}\Leftrightarrow f(\Gamma) \subseteq f(\varphi)$, para toda f.

d: **($\Rightarrow$):** sejam $\Gamma \vDash_{\mathbb{P}} \varphi$ e $\Gamma \nvDash_{\mathcal{V}} \varphi$. Então existe $\mathcal{V}$ tal que, para cada $\varphi_i \in \Gamma$, $\mathcal{V}(\varphi_i) = V$ e $\mathcal{V}(\varphi) = F \Rightarrow_{\text{Teo5.4.6a}}\Rightarrow$ existe f_j tal que $P(f_j)(\varphi_i) = 1$, para cada $\varphi_i \in \Gamma$ e existe f_k tal que $P(f_k)(\varphi) = 0$. Neste caso, pode-se mostrar que existe f tal que $P(f)(\varphi_i) = 1$, para cada $\varphi_i \in \Gamma$ e $P(f)(\varphi) = 0 \Rightarrow P(f)(\Gamma) > P(f)(\varphi) \Rightarrow \Gamma \nvDash_{\mathbb{P}} \varphi$, contrariando a hipótese inicial.

($\Leftarrow$): sejam $\Gamma \vDash_{\mathcal{V}} \varphi$ e $\Gamma \nvDash_{\mathbb{P}} \varphi$. Então existe f tal que $P(f)(\Gamma) > P(f)(\varphi)$. Seja f_j tal que $P(f_j)(\Gamma) = 1$ e $P(f_j)(\varphi) = 0 \Rightarrow$ para cada $\varphi_i \in \Gamma$, $P(f_j)(\varphi_i) = 1$ e $P(f_j)(\varphi) = 0 \Rightarrow_{\text{Teo5.4.6a}}\Rightarrow$

existe $\mathcal{V}$ tal que, para cada $\varphi_i \in \Gamma$, $\mathcal{V}(\varphi_i) = V$ e $\mathcal{V}(\varphi) = F \Rightarrow \Gamma \nvDash_{\mathcal{V}} \varphi$, contrariando a hipótese inicial.

e: $\Gamma \vDash_{\mathbb{P}} \varphi \Leftrightarrow_{\text{Teo6.3.7d}} \Leftrightarrow \Gamma \vDash_{\mathcal{V}} \varphi \Leftrightarrow \Gamma \vdash \varphi$. ■

Apesar de assumirmos *LSC* como sistema formal lógico subjacente, as propriedades acima são gerais, valem para qualquer sistema que adote as definições envolvidas. Ou seja, as propriedades são dos conceitos e independem do sistema subjacente.

O Teorema 6.3.7a explicita a relação entre duas definições de $S_{\mathbb{P}}$: o conjunto de fórmulas que são consequência lógica probabilística do vazio é igual ao conjunto de fórmulas probabilisticamente válidas. Embora se refiram aos mesmos elementos, as duas definições possuem conteúdos distintos. Enquanto a primeira é uma relação entre o conjunto vazio de fórmulas e uma fórmula, a segunda é uma propriedade das fórmulas. Não faz sentido, por exemplo, dizer que $\vDash_{\mathbb{P}} \varphi$ possui um valor probabilidade, ao contrário de $\top_{\mathbb{P}}$, que se caracteriza exatamente por possuir determinado valor probabilidade.

O Teorema 6.3.7b explicita que uma fórmula é consequência lógica probabilística de um dado conjunto de fórmulas sse essa fórmula é probabilisticamente certa em todo modelo do conjunto dado. Esse resultado mostra que a consequência lógica probabilística pode ser definida a partir da noção de modelo, como feito usualmente na consequência lógica veritativo-funcional: uma fórmula φ é consequência lógica de um conjunto Γ de fórmulas sse φ é válida em todo modelo de Γ.

Apesar dessa correspondência entre as definições de consequência lógica probabilística e Veritativo-Funcional, podemos apontar uma diferença oriunda destas duas definições. Na perspectiva veritativo-funcional, pode-se mostrar que: $\Gamma \vDash_{\mathcal{V}} \varphi$, para $\Gamma = \{\varphi_1, ..., \varphi_n\} \Rightarrow \mathcal{V}(\varphi_i) \leq \mathcal{V}(\varphi)$, para algum $\varphi_i \in \Gamma$. Considerando o valor verdade "V" como "1" e o valor de verdade "F" como "0", isso é bastante óbvio. Se $\mathcal{V}(\varphi_i) > \mathcal{V}(\varphi)$, para toda $\varphi_i \in \Gamma$, então a valoração $\mathcal{V}$ seria modelo de Γ mas não seria modelo de φ, ou seja, φ não seria válida em $\mathcal{V}$. Consequentemente, $\Gamma \nvDash_{\mathcal{V}} \varphi$.

O resultado mostrado acima não vale para a consequência lógica probabilística: $\Gamma \vDash_{\mathbb{P}} \varphi$, para $\Gamma = \{\varphi_1, ..., \varphi_n\} \nRightarrow P(\varphi_i) \leq P(\varphi)$, para algum $\varphi_i \in \Gamma$. O seguinte caso serve para ilustrar: consideremos a situação do jogo de dados, tal como no Exemplo 5.3.3, em que $P(A_1) = ½$, $P(A_3) = 2/3$, $P(A_1 \rightarrow \neg A_3) = 5/6$, $P(\neg A_3) = 1/3$. Neste caso, $\{A_1, A_1 \rightarrow \neg A_3\}$

$\vDash_{\mathbb{P}} \neg A_3$, embora a probabilidade de cada uma das duas premissas seja maior do que a probabilidade da conclusão.

Essa característica própria da consequência lógica probabilística ocorre devido à Definição 6.3.1 estabelecer a probabilidade do conjunto de premissas como definidor da consequência. Do Teorema 5.4.11a, temos que, para $\Gamma = \{\varphi_1, ..., \varphi_n\}$, $P(\Gamma) \leq P(\varphi_i)$, para toda $\varphi_i \in \Gamma$. O conjunto Γ é pelo menos tão preciso quanto cada φ_i, o que o torna no mínimo, tão arriscado quanto φ_i.

Tal diferença entre as duas definições de consequência lógica desaparece quando consideramos apenas os valores probabilidade extremos 0 e 1. Neste caso, $\Gamma \vDash_{\mathbb{P}} \varphi$, para $\Gamma = \{\varphi_1, ..., \varphi_n\} \Rightarrow P(\varphi_i) \leq P(\varphi)$, para algum $\varphi_i \in \Gamma$. Isto porque $P(\varphi_i) > P(\varphi)$, para toda $\varphi_i \in \Gamma \Rightarrow P(\varphi_i) = 1$, para toda $\varphi_i \in \Gamma$. Além disso, nas condições atuais, conclui-se que $P(\varphi) = 0$. Assim, pelo Teorema 5.3.6f, $P(\Gamma) = 1$ e $P(\varphi) = 0$. Logo, $\Gamma \nvDash_{\mathbb{P}} \varphi$. Com isso, concluímos, outra vez, que, quando se trata unicamente de valores extremos, $S_{\mathbb{P}}$ e $S_{\mathcal{V}}$ funcionam do mesmo modo.

O Teorema 6.3.7c expressa uma característica fundamental da consequência lógica probabilística: a conclusão está garantida, a partir da garantia das premissas. Quando o valor probabilidade do conjunto de premissas é sempre menor do que o valor probabilidade da conclusão, o evento associado a ele está contido, em toda situação, ao evento associado à conclusão. Assim, obviamente, acontecendo um dos elementos daquele, garantimos o acontecimento de um dos elementos deste.

Embora as premissas garantam a conclusão em uma consequência lógica, há possibilidade de ocorrência da conclusão sem que as premissas ocorram. Para ilustrar isso, observemos o argumento do seguinte exemplo.

Exemplo 6.3.8

Se cai número ímpar, então cai número primo.

Cai número ímpar.

Cai número primo.

Neste caso, considerando o jogo de um dado não viciado como situação, $f(\Gamma) = \{1,3,5\} \subseteq f(\varphi) = \{1,2,3,5\}$. Se o número que caiu for primo, ou seja, o evento correspondente à conclusão ocorre, não há garantia absoluta de que ele tenha sido ímpar, ou seja, que o evento correspondente à segunda premissa tenha ocorrido.

No entanto, o argumento acima evidencia que, se ocorre o evento associado ao conjunto de premissas, também ocorre o evento associado à conclusão. Se isto não acontecesse, então teríamos uma contradição: algum elemento de $f(\Gamma)$ teria e não teria ocorrido ao mesmo tempo.

O Teorema 6.3.7d expressa que a consequência lógica probabilística e a consequência lógica veritativo-funcional são satisfeitos pelo mesmo conjunto de objetos. Discutiremos nas Considerações Finais os desdobramentos destes dois últimos resultados para a consequência lógica veritativo-funcional.

O Teorema 6.3.7e é o *Teorema da Completude para LSC a partir da consequência lógica probabilística*. Um caso particular do Teorema 6.3.7e é quando consideramos o conjunto de premissas vazio: $\vDash_{\mathbb{P}} \varphi \Leftrightarrow \vdash \varphi$. Ele pode ser provado sem recorrermos ao caso geral: $\vDash_{\mathbb{P}} \varphi \Leftrightarrow_{\text{Teo6.3.7a}}\Leftrightarrow \varphi$ é $\top_{\mathbb{P}} \Leftrightarrow_{\text{Teo5.4.6c}}\Leftrightarrow \vdash \varphi$. Este caso específico mostra que os conceitos de consequência lógica probabilística do vazio e de consequência lógica sintática do vazio, ou seja, dos teoremas de *L*, são satisfeitos pelas mesmas fórmulas de *L*. Juntamente com o Teorema 6.3.7a, este último resultado explicita que os conceitos de consequência lógica probabilística do vazio, consequência lógica sintática do vazio e validade probabilística se referem às mesmas fórmulas de *L* na *LSC*.

De acordo com as Definições 6.3.1 e 6.3.4, a rigor o conjunto de premissas em uma consequência lógica probabilística é sempre finito. No entanto, do Teorema 6.3.7e, podemos, de alguma forma, eliminar essa restrição do conjunto de premissas, garantindo, inclusive, o *Teorema da Compacidade* para a consequência lógica probabilística.

No próximo teorema mostramos outras propriedades gerais da consequência lógica probabilística. Todos os resultados também podem ser mostrados para a consequência lógica veritativo-funcional. Escreveremos "Γ, φ" como uma abreviação para "$\Gamma \cup \{\varphi\}$".

Teorema 6.3.9

a: $\Gamma, \varphi \vDash_{\mathbb{P}} \psi \Leftrightarrow \Gamma \vDash_{\mathbb{P}} \varphi \rightarrow \psi$.

b: $\vDash_{\mathbb{P}} \varphi \Leftrightarrow \Gamma \vDash_{\mathbb{P}} \varphi$, para todo Γ.

c: $P(\Gamma) = 1$, para toda f, e $\Gamma \vDash_{\mathbb{P}} \varphi \Leftrightarrow \vDash_{\mathbb{P}} \varphi$.

d: $\varphi \vDash_{\mathbb{P}} \psi$ e $\psi \vDash_{\mathbb{P}} \gamma \Rightarrow \varphi \vDash_{\mathbb{P}} \gamma$.

e: $\vDash_{\mathbb{P}} \varphi$ e $\vDash_{\mathbb{P}} \psi \Rightarrow \varphi \equiv \psi$.

f: $\varphi \equiv \psi \Leftrightarrow\ \vDash_{\mathbb{P}} \varphi \leftrightarrow \psi$.

Demonstração: Seja $\Gamma = \{\varphi_1, ..., \varphi_n\}$.

a: Por indução no número n de fórmulas em Γ.

Base: n = 0, ou seja, Γ é vazio. Temos de mostrar que $\varphi \vDash_{\mathbb{P}} \psi \Leftrightarrow\ \vDash_{\mathbb{P}} \varphi \to \psi$: $\varphi \vDash_{\mathbb{P}} \psi \Leftrightarrow_{Def6.3.4}\Leftrightarrow P(\varphi) \leq P(\psi)$, para toda $f \Leftrightarrow_{Teo5.4.2b}\Leftrightarrow \varphi \to \psi$ é $\top_{\mathbb{P}} \Leftrightarrow_{Teo6.3.9b}\Leftrightarrow\ \vDash_{\mathbb{P}} \varphi \to \psi$.

Passo Indutivo: assumamos que o resultado vale para todo i < n.

Seja $\Gamma_i = \{\varphi_1, ..., \varphi_i\} \subset \Gamma$. Daí, $P(\Gamma_i) \leq_{Teo5.4.11a}\leq P(\Gamma) \Rightarrow P(\Gamma_i \wedge \varphi) \leq P(\Gamma \wedge \varphi)$. Assim, $\Gamma_i, \varphi \vDash_{\mathbb{P}} \psi \Rightarrow_{Def6.3.4}\Rightarrow \Gamma, \varphi \vDash_{\mathbb{P}} \psi$ e $\Gamma_i \vDash_{\mathbb{P}} \varphi \to \psi \Rightarrow_{Def6.3.4}\Rightarrow \Gamma \vDash_{\mathbb{P}} \varphi \to \psi$. Pela hipótese da indução, tem-se que $\Gamma_i, \varphi \vDash_{\mathbb{P}} \psi \Leftrightarrow \Gamma_i \vDash_{\mathbb{P}} \varphi \to \psi$. Logo, $\Gamma, \varphi \vDash_{\mathbb{P}} \psi \Leftrightarrow \Gamma \vDash_{\mathbb{P}} \varphi \to \psi$.

b: $\vDash_{\mathbb{P}} \varphi \Leftrightarrow_{Teo6.3.7a}\Leftrightarrow \varphi$ é $\top_{\mathbb{P}} \Leftrightarrow_{Def5.4.1b}\Leftrightarrow P(\varphi) = 1$, para toda $f \Leftrightarrow P(\Gamma) \leq P(\varphi)$, para toda f e para todo $\Gamma \Leftrightarrow_{Def6.3.4}\Leftrightarrow \Gamma \vDash_{\mathbb{P}} \varphi$, para todo Γ.

c: **(⇒):** $P(\Gamma) = 1$, para toda f, e $\Gamma \vDash_{\mathbb{P}} \varphi \Rightarrow_{Def6.3.4}\Rightarrow P(\Gamma) \leq P(\varphi)$, para toda $f \Rightarrow P(\varphi) = 1$ para toda $f \Rightarrow_{Def5.4.1b}\Rightarrow \varphi$ é $\top_{\mathbb{P}} \Rightarrow_{Teo6.3.7a}\Rightarrow\ \vDash_{\mathbb{P}} \varphi$.

(⇐): $\vDash_{\mathbb{P}} \varphi \Rightarrow_{Teo6.3.9b}\Rightarrow \Gamma \vDash_{\mathbb{P}} \varphi$, para todo Γ.

d: $\varphi \vDash_{\mathbb{P}} \psi$ e $\psi \vDash_{\mathbb{P}} \gamma \Rightarrow_{Def6.3.4}\Rightarrow P(\varphi) \leq P(\psi) \leq P(\gamma)$, para toda $f \Rightarrow P(\varphi) \leq P(\gamma)$, para toda f $\Rightarrow_{Def6.3.4}\Rightarrow \varphi \vDash_{\mathbb{P}} \gamma$.

e: $\vDash_{\mathbb{P}} \varphi$ e $\vDash_{\mathbb{P}} \psi \Leftrightarrow_{Teo6.3.7a}\Leftrightarrow \varphi$ é $\top_{\mathbb{P}}$ e ψ é $\top_{\mathbb{P}} \Rightarrow P(\varphi) = P(\psi) = 1$, para toda $f \Rightarrow_{Def5.4.8}\Rightarrow \varphi \equiv \psi$.

f: $\varphi \equiv \psi \Leftrightarrow_{Def5.4.9b}\Leftrightarrow \varphi \leftrightarrow \psi$ é $\top_{\mathbb{P}}\Leftrightarrow_{Teo6.3.7a}\Leftrightarrow\ \vDash_{\mathbb{P}} \varphi \leftrightarrow \psi$. ■

Uma das partes do primeiro item acima é uma propriedade semântica correspondente ao Teorema da Dedução para *LSC*: $\Gamma, \varphi \vdash \psi \Rightarrow \Gamma \vdash \varphi \to \psi$. Em particular, temos que: $\varphi \vDash_{\mathbb{P}} \psi \Leftrightarrow\ \vDash_{\mathbb{P}} \varphi \to \psi$, ou seja, ψ é consequência lógica probabilística de φ sse $\varphi \to \psi$ é probabilisticamente válida.

O segundo item explicita que as fórmulas que são consequências lógicas probabilísticas do vazio também o são de qualquer conjunto de fórmulas. De algum modo, podemos dizer que as fórmulas que se auto-sustentam são também sustentadas por qualquer conjunto de fórmulas. Uma das partes deste resultado, qual seja, "$\vDash_{\mathbb{P}} \varphi \Rightarrow \Gamma \vDash_{\mathbb{P}} \varphi$, para todo Γ”, é um caso particular da propriedade da monotonicidade da consequência lógica probabilística, a ser provada no próximo teorema.

O terceiro item esclarece que as fórmulas que são consequência do vazio também o são de qualquer conjunto de fórmulas probabilisticamente válidas. Em especial, elas são as únicas fórmulas que são consequências de conjuntos constituídos unicamente de axiomas de *LSC*.

O quarto item do teorema acima explicita que a consequência lógica probabilística satisfaz a propriedade da transitividade. A recíproca deste resultado não vale: $\varphi \vDash_{\mathbb{P}} \gamma \not\Rightarrow \varphi \vDash_{\mathbb{P}} \psi$ e $\psi \vDash_{\mathbb{P}} \gamma$. Isto pode ser assim mostrado: do Exemplo 5.3.3, temos que: $A_1 \wedge A_2 \vDash_{\mathbb{P}} A_1$, mas não vale que $A_1 \wedge A_2 \vDash_{\mathbb{P}} A_3$ e $A_3 \vDash_{\mathbb{P}} A_1$, dado que $A_1 \wedge A_2 \nvDash_{\mathbb{P}} A_3$.

O quinto item expressa que, se duas fórmulas são consequência probabilística do vazio, então são probabilisticamente equivalentes. Este resultado também ilustra que a consequência lógica probabilística não captura a noção de conteúdo informacional. Temos, por exemplo, que $\vDash_{\mathbb{P}} \varphi \vee \neg\varphi$ e $\vDash_{\mathbb{P}} \neg(\varphi \wedge \neg\varphi)$. Mas, como já discutido em comentários anteriores, especialmente aquele referente ao Teorema 6.2.4c, embora essas duas fórmulas possuam sempre o mesmo valor probabilidade, nem sempre querem dizer a mesma coisa.

A recíproca do resultado comentado acima não vale: $\varphi \equiv \psi \not\Rightarrow \vDash_{\mathbb{P}} \varphi$ e $\vDash_{\mathbb{P}} \psi$. Por exemplo: $\varphi \rightarrow \psi \equiv \neg\varphi \vee \psi$, mas não ocorre que: $\vDash_{\mathbb{P}} \varphi \rightarrow \psi$ e $\vDash_{\mathbb{P}} \neg\varphi \vee \psi$, dado que nenhuma das duas fórmulas é probabilisticamente válida.

Por fim, o sexto item do teorema em questão expressa que duas fórmulas probabilisticamente equivalentes constituem uma fórmula biimplicativa que é consequência lógica probabilística do vazio, ou seja, probabilisticamente válida.

Para terminarmos a análise da consequência lógica probabilística, mostramos a seguir que ela é uma consequência lógica Tarskiana. Ou seja, ele satisfaz a propriedades da reflexividade, monotonicidade e transitividade, tal como apresentados na Definição 1.3.1.

Teorema 6.3.10

a: $\varphi \in \Gamma \Rightarrow \Gamma \vDash_{\mathbb{P}} \varphi$.

b: $\Gamma \subseteq \Delta$ e $\Gamma \vDash_{\mathbb{P}} \varphi \Rightarrow \Delta \vDash_{\mathbb{P}} \varphi$.

c: $\Delta \vDash_{\mathbb{P}} \psi$, para cada $\psi \in \Gamma$, e $\Gamma \vDash_{\mathbb{P}} \varphi \Rightarrow \Delta \vDash_{\mathbb{P}} \varphi$.

Demonstração: Seja f uma dada situação qualquer.

a: $\Gamma = \{\varphi, ..., \varphi_n\}$, para $n \geq 0 \Rightarrow_{Def6.3.1}\Rightarrow P(\Gamma) = P(\varphi \wedge ... \wedge \varphi_n) \Rightarrow_{Teo5.4.11a}\Rightarrow P(\varphi \wedge ... \wedge \varphi_n) \leq P(\varphi) \Rightarrow_{Def6.3.4}\Rightarrow \Gamma \vDash_{\mathbb{P}} \varphi$.

b: Sejam $\Gamma= \{\varphi_1, ..., \varphi_n\}$, $\Delta = \{\varphi_1, ..., \varphi_n, \varphi_{n+0}, ..., \varphi_{n+m}\}$, para $n, m \in \mathbb{N} \Rightarrow_{Teo5.4.11a}\Rightarrow P(\Gamma) \leq P(\Delta)$. Assim, $\Gamma \vDash_{\mathbb{P}} \varphi \Rightarrow_{Def6.3.1}\Rightarrow P(\Gamma) \leq P(\varphi) \Rightarrow P(\Delta) \leq P(\varphi) \Rightarrow \Delta \vDash_{\mathbb{P}} \varphi$.

c: $\Delta \vDash_{\mathbb{P}} \psi$, para cada $\psi \in \Gamma$, e $\Gamma \vDash_{\mathbb{P}} \varphi \Rightarrow_{Teo6.3.7c}\Rightarrow f(\Delta) \subseteq f(\Gamma) \subseteq f(\varphi)$, para toda $f \Rightarrow f(\Delta) \subseteq f(\varphi) \Rightarrow_{Teo6.3.7c}\Rightarrow \Delta \vDash_{\mathbb{P}} \varphi$. ■

A reflexividade é satisfeita pela consequência lógica probabilística pelo fato de que a probabilidade de uma fórmula ser sempre menor ou igual do que o valor probabilidade de qualquer conjunto que a contém como elemento.

A monotonicidade é satisfeita pela consequência lógica probabilística porque o acréscimo de fórmulas em um conjunto faz o valor probabilidade do novo conjunto ser menor ou igual ao valor do conjunto original.

Devido às definições relacionadas à consequência lógica probabilística, encontramos na propriedade da Transitividade uma especificidade, não encontrada, por exemplo, na consequência lógica veritativo-funcional. Do Teorema 5.4.11a obtemos que $P((\varphi_1 \wedge ... \wedge \varphi_n) \wedge \psi) \leq \psi$. Assim, se $\Gamma = \{\psi_1, ..., \psi_n\}$, temos que $P(\Gamma) \leq \psi_i$, para todo $\psi_i \in \Gamma$. Desse modo, embora $\Gamma \vDash_{\mathbb{P}} \varphi$, poderíamos supor a existência de uma situação em que $P(\Delta) \leq \psi_i$, para cada $\psi_i \in \Gamma$, mas $P(\Delta) > P(\Gamma)$, podendo resultar que $\Delta \nvDash_{\mathbb{P}} \varphi$. No entanto, isso jamais acontece, como mostrado no teorema acima.

Com isto, terminamos a análise da consequência lógica probabilística. Dentre seus principais resultados, concluímos que, embora seja satisfeita pelos mesmos objetos que a consequência lógica veritativo-funcional e que seja uma consequência lógica Tarskiana, ela apresenta alguns comportamentos próprios.

No que se segue tratamos da consequência lógica informacional, mostrando alguns resultados próprios atribuídos a ela.

4 Consequência lógica informacional

Nesta seção introduzimos a definição de consequência lógica informacional e mostramos algumas propriedades específicas desta definição. Apontaremos, particularmente, que ela não é equivalente à consequência lógica probabilística. Provamos, ainda, que nem toda fórmula é consequência lógica informacional de um conjunto contraditório de fórmulas, que a consequência lógica informacional não é uma consequência lógica Tarskiana e que a Lógica subjacente a ela não é a Lógica Clássica.

Iniciamos definindo a quantidade de informação em um conjunto de fórmulas de uma linguagem de *LSC*, ponto de partida para a definição de consequência lógica informacional.

Definição 6.4.1 (Quantidade de informação em um conjunto de fórmulas)

Seja $\Gamma = \{\varphi_1, ..., \varphi_n\} \subseteq$ Form(L). O *valor informacional de Γ* ou *a quantidade de informação em Γ, segundo* $f(\Sigma)$, denotado por "$I^*(f(\Sigma))(\Gamma)$", é definido pela seguinte equação:

$$I^*(f(\Sigma))(\Gamma) = -\log_2 P^*(f(\Sigma))(\Gamma).$$

Quando não gerar ambiguidade ou imprecisão, escreveremos $I(\Gamma)$ em vez de $I^*(f(\Sigma))(\Gamma)$. Observe-se que, embora o domínio desta função seja diferente do domínio da função informação sobre fórmulas, o seu conjunto imagem é o mesmo. Ou seja, $I^*(f(\Sigma))$: $\wp(\text{Form}(L)) \to \mathbb{Q}_+$.

Teorema 6.4.2

Γ é contraditório $\Rightarrow I(\Gamma) = 0$, para toda f.

Demonstração

Γ é contraditório $\Rightarrow_{\text{Teo6.3.3c}}\Rightarrow P(\Gamma) = 0$, para toda $f \Rightarrow_{\text{Def6.4.1}}\Rightarrow I(\Gamma) = 0$, para toda f. ∎

A recíproca deste teorema não pode ser provada nas condições atuais, ou seja, $I(\Gamma) = 0$, para toda $f \not\Rightarrow \Gamma$ é contraditório, como ilustra o seguinte exemplo: $\Gamma = \{\top_{\mathbb{P}}\} \Rightarrow I(\Gamma) = 0$ para toda f, embora Γ não seja contraditório.

A partir deste teorema mostraremos algumas diferenças entre a consequência lógica probabilística e a consequência lógica informacional, da qual passamos a tratar em seguida.

Antes de introduzirmos a definição da consequência lógica informacional, elencamos algumas ideias intuitivas a serem adotadas como ponto de partida e referência para tal definição.

Em uma consequência lógica, ou em uma dedução ou um raciocínio válido, a informação presente na conclusão, está, de algum modo, presente nas premissas. Observemos os seguintes silogismos, representados graficamente através dos clássicos Diagramas de Venn.

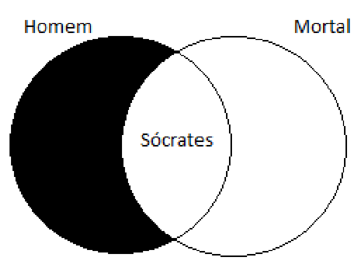
Exemplo 6.4.3

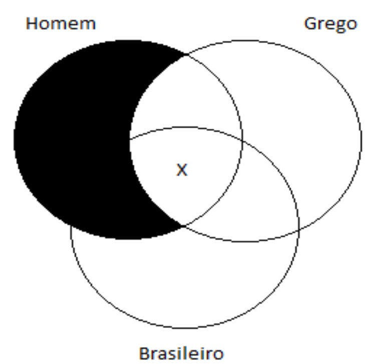
a:

Todo homem é mortal.
Sócrates é homem.
Sócrates é mortal.

b:

Todo homem é grego.
Algum homem é brasileiro.
Algum brasileiro é grego.





O gráfico acima ilustra que as premissas do primeiro argumento já garantem que Sócrates é mortal, conforme explicitado pela conclusão. O conjunto de premissas, inclusive, é mais preciso, mais informativo do que a conclusão. Conjuntamente, dizem que Sócrates, além de ser mortal, também é homem, enquanto a conclusão, isoladamente, nada informa sobre a relação entre Sócrates e homem. Ela garante apenas que Sócrates é mortal, podendo ou não ser homem.

Por ser menos informativa (mais provável), a conclusão do exemplo em questão poderia ser satisfeita, ou seja, o evento correspondente a ela em uma dada situação teria ocorrido, ainda que o conjunto de premissas não o fosse. Este seria o caso em que Sócrates fosse mortal sem ser um humano, circunstância em que a primeira premissa não ocorreria.

O gráfico referente ao segundo argumento do exemplo acima mostra algo semelhante ao primeiro. Nele, fica garantido, a partir das premissas, que algum brasileiro é grego. Ou seja, as premissas garantem a conclusão. O conjunto de premissas é mais informativo do que a conclusão, pois diz ainda que os referidos elementos que são brasileiros e gregos também são homens. A conclusão, por sua vez, determina a existência de indivíduos brasileiros e gregos, sem nada dizer a respeito da sua relação com o conjunto dos homens. Assim, a conclusão está garantida quando o conjunto de premissas o estiver.

Estes dois argumentos, que parecem ser intuitivamente válidos, se caracterizam por apresentar menos informação na conclusão daquela apresentada pelo seu conjunto de premissas.

Os dois próximos argumentos ilustram casos em que a informação da conclusão é mais ampla do que a informação contida no conjunto de premissas, o que os torna argumentos inválidos. Intuitivamente, percebemos que a conclusão não é consequência lógica das premissas.

Exemplo 6.3.4

a:

João é paulista e brasileiro.

Maria é paulista e brasileira.

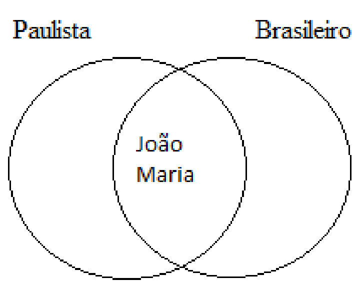
Todo paulista é brasileiro.

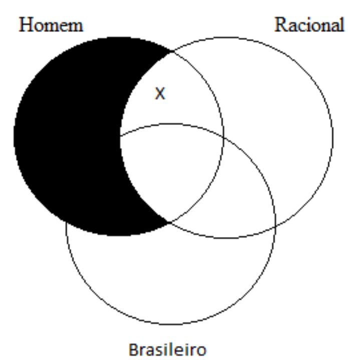
b:

Todo homem é racional.

Algum homem não é brasileiro.

Algum brasileiro é racional.





No primeiro argumento acima há informações na conclusão que extrapolam a informação presente nas premissas. A conclusão diz que os outros componentes do conjunto dos paulistas, quando houver, também são brasileiros. Assim, a informação presente nas premissas não garante a informação presente na conclusão.

No segundo argumento, não está garantido, a partir das premissas, a existência de objetos que são brasileiros e racionais ao mesmo tempo. Não há, nas premissas, essa informação, presente na conclusão. Como esta contém informação inexistente naquelas, não é garantida pelas premissas.

Embora os quatro argumentos acima não possam ser adequadamente expressos na linguagem de *LSC*, eles servem de base intuitiva para a análise informacional da consequência lógica: uma fórmula é consequência lógica informacional de um dado conjunto de premissas sse a informação presente na conclusão não extrapola a informação presente nas premissas.

Ressaltamos que a noção de informação a ser utilizada é a quantificacional, tal como exposta no quarto capítulo deste trabalho, desenvolvida por pensadores como Shannon (1949) e adotada na TMC. Não consideraremos os elementos semânticos da informação ou o conteúdo informacional das sentenças envolvidas. A ideia básica, portanto, é que uma fórmula é consequência lógica informacional de um dado conjunto de fórmulas sse a quantidade de informação presente na conclusão jamais pode ser maior do que a quantidade de informação das premissas, conforme a próxima definição.

Definição 6.4.5 (Consequência lógica informacional)

Uma fórmula φ é *consequência lógica informacional* de um conjunto Γ de sentenças, denotado por "$\Gamma \Vdash \varphi$", sse, para toda $f(\Sigma)$, $I(f(\Sigma))(\Gamma) \geq I(f(\Sigma))(\varphi)$.

Quando $\Gamma = \varnothing$, em vez de "$\Gamma \Vdash \varphi$" escreveremos simplesmente "$\Vdash \varphi$". "$\Gamma \nVdash \varphi$" denota que φ não é consequência lógica informacional de Γ. Denotaremos o conjunto de consequências informacional de Γ por $\Vdash(\Gamma)$, ou seja: $\Vdash(\Gamma) = \{\varphi \mid \Gamma \Vdash \varphi\}$.

A seguir apresentamos alguns pares de fórmulas tal que uma é consequência lógica informacional da outra. Eles poderiam ser provados como um corolário dos próximos teoremas. Em discussões futuras, serão utilizados para analisar a adequação intuitiva da definição de consequência lógica informacional.

Teorema 6.4.6

a: $\varphi \to \psi \Vdash\!\models \neg\varphi \vee \psi$.

b: $\varphi \vee \psi \Vdash\!\models \neg(\neg\varphi \wedge \neg\psi)$.

c: $\varphi \wedge \psi \Vdash\!\models \neg(\neg\varphi \vee \neg\psi)$.

d: $\varphi \wedge \psi \Vdash\!\models \psi \wedge \varphi$.

e: $\varphi \vee \psi \Vdash\!\models \psi \vee \varphi$.

f: $\neg\neg\varphi \Vdash\!\models \varphi$.

g: $\varphi \wedge \varphi \Vdash\!\models \varphi$.

h: $\varphi \vee \varphi \Vdash\!\models \varphi$.

i: $\top_{\mathbb{P}} \wedge \varphi \Vdash\!\models \varphi$.

j: $\top_{\mathbb{P}} \vee \varphi \Vdash\!\models \top_{\mathbb{P}}$.

k: $\perp_{\mathbb{P}} \wedge \varphi \Vdash\!\models \perp_{\mathbb{P}}$.

l: $\perp_{\mathbb{P}} \vee \varphi \Vdash\!\models \varphi$.

m: $\top_{\mathbb{P}} \Vdash\!\models \perp_{\mathbb{P}}$.

n: $\varphi \Vdash\!\models \varphi$.

Demonstração

São demonstrados pelo Teorema 6.2.5 e Definição 6.4.5. ■

Como as fórmulas em cada um dos pares são informacionalmente iguais, a recíproca de cada um dos itens do teorema acima também vale. Este teorema é correspondente ao Teorema 6.3.5, onde mostramos que estes mesmos pares de fórmulas satisfazem a noção de consequência lógica probabilística.

Destacamos que os três primeiros casos do teorema acima mostram que as definições de um conectivo a partir de outros, como na Definição 2.2.6, são mantidas na consequência lógica informacional. O primeiro item explicita que a noção de implicação pressuposta aqui é a da implicação material. O quarto e o sexto itens mostram, respectivamente, que a Lógica subjacente à consequência lógica informacional não é uma lógica temporal nem uma lógica intuicionista.

O próximo teorema explicita uma relação entre as fórmulas que são consequências lógicas do vazio com a sua quantidade de informação, além de uma relação entres as definições de consequência lógica probabilística e informacional.

Teorema 6.4.7

a: $\Vdash\!\models \varphi \Leftrightarrow I(\varphi) = 0$, para toda *f*.

b: $\Vdash\!\models \varphi \Leftrightarrow \Gamma \Vdash\!\models \varphi$, para todo Γ.

c: $I(\Gamma) = 0$, para toda *f*, e $\Gamma \Vdash\!\models \varphi \Rightarrow \Vdash\!\models \varphi$.

d: $\models_{\mathbb{P}} \varphi \Rightarrow \Vdash\!\models \varphi$.

Demonstração

a: **(⇒):** $\Vdash \varphi \Rightarrow_{Def6.4.5}\Rightarrow I(\Gamma) = 0 \geq I(\varphi)$, para toda $f \Rightarrow I(\varphi) = 0$, para toda f.

(⇐): $I(\varphi) = 0$, para toda $f \Rightarrow$ para toda f, $I(\varnothing) \geq I(\varphi) \Rightarrow_{Def6.4.5}\Rightarrow \Vdash\varphi$.

b: **(⇒):** $\Vdash\varphi \Rightarrow_{Teo6.4.7a}\Rightarrow I(\varphi) = 0$, para toda $f \Rightarrow I(\Gamma) \geq I(\varphi)$, para todo Γ e toda $f \Rightarrow_{Def6.4.5}\Rightarrow$ $\Gamma \Vdash \varphi$, para todo Γ.

(⇐): $\Gamma \Vdash \varphi$, para todo $\Gamma \Rightarrow \varnothing \Vdash \varphi \Rightarrow_{Def6.3.4}\Rightarrow \Vdash \varphi$.

c: $I(\Gamma) = 0$, para toda f, e $\Gamma \Vdash \varphi \Rightarrow_{Def6.3.4}\Rightarrow I(\Gamma) \geq I(\varphi)$, para toda $f \Rightarrow I(\varphi) = 0$, para toda $f \Rightarrow_{Teo6.4.7a}\Rightarrow \Vdash \varphi$.

d: $\vDash_{\mathbb{P}} \varphi \Rightarrow_{Teo6.3.7a}\Rightarrow \varphi$ é $\top_{\mathbb{P}} \Rightarrow_{Def.5.38b}\Rightarrow P(\varphi) = 1$, para toda $f \Rightarrow_{Def6.2.1}\Rightarrow I(\varphi) = 0$, para toda $f \Rightarrow_{Teo6.4.7a}\Rightarrow \Vdash \varphi$. ■

O primeiro item do teorema acima explicita que as fórmulas que se auto sustentam não possuem informação. Como a informação da conclusão deve ser sempre menor ou igual à informação do conjunto de premissas, obviamente as conclusões que são consequências lógicas informacionais do vazio também devem ser informacionalmente vazias.

O resultado em questão mostra que, para este caso, a consequência lógica informacional captura a noção de informação da TMC. As fórmulas que são informacionalmente vazias possuem probabilidade máxima ou mínima em toda situação. Elas não reduzem incerteza, por se referirem a algo absolutamente certo ou a algo absolutamente impossível. Estas fórmulas, de acordo com o teorema acima, são aquelas que se auto sustentam. Se forem probabilisticamente válidas, não precisam de nada além delas próprias para garantir a ocorrência do evento associado a ela. Se forem probabilisticamente contraditórias, garantem, por si só, a não ocorrência do evento associado a ela por uma dada situação qualquer.

O Teorema 6.4.7c, aparentemente óbvio, expressa uma propriedade da consequência lógica informacional que não pode ser mostrada para as noções de consequência lógica veritativo-funcional e probabilística. Quando $I(\Gamma) = 0$, o conjunto Γ poder vazio, pode ser contraditório (possuir probabilidade zero em toda situação) ou pode ser probabilisticamente certo (possuir probabilidade um em toda situação). Assim, ao averiguar que uma fórmula é consequência lógica informacional de algum destes conjuntos, também mostramos que ele é consequência lógica informacional do conjunto vazio (e também dos outros dois).

O resultado acima não vale nas perspectivas probabilística e veritativo-funcional. Isso porque, em ambas as perspectivas, o fato de uma fórmula ser consequência lógica de um conjunto contraditório de premissas, não garante que ela seja consequência lógica do vazio. Esta diferença da versão informacional com as outras duas se origina, outra vez, no fato de a perspectiva informacional não distinguir tautologias de contradições, considerando ambas informacionalmente vazias.

Dos três primeiros itens do teorema em questão podemos obter que as fórmulas informativas são aquelas que são consequência lógica de um conjunto de fórmulas que também é informativo, ou seja, $\Gamma \Vdash \varphi$ e $I(\varphi) \neq 0$, para alguma $f \Rightarrow I(\Gamma) \neq 0$, para alguma f. Isso implica que a satisfação da conclusão é garantida a partir da satisfação de um conjunto informativo de premissas. Ela não se garante por si só. Como o conjunto de premissas não é absolutamente certo ou impossível, quando ele não for satisfeito, a conclusão também pode não ser satisfeita e quando ele o for, também o será a conclusão.

O Teorema 6.4.7d explicita que as consequências lógicas probabilísticas do vazio também são consequências lógicas informacionais do vazio. A recíproca, no entanto, não pode ser mostrada: $\Vdash \varphi \not\Rightarrow \vDash_{\mathbb{P}} \varphi$. O seguinte exemplo serve para mostrar essa invalidade: $\Vdash \perp_{\mathbb{P}}$, mas $\nvDash_{\mathbb{P}} \perp_{\mathbb{P}}$. Esse resultado é muito expressivo. Ele mostra que a consequência lógica informacional não é equivalente à consequência lógica probabilística e, consequentemente, à consequência lógica veritativo-funcional.

O Teorema 6.4.7d, no entanto, não pode ser mostrado quando os conjuntos de hipóteses são informativos. Ou seja, o resultado não vale para o caso geral: $\Gamma \vDash_{\mathbb{P}} \varphi \not\Rightarrow \Gamma \Vdash \varphi$. No Teorema 6.3.6 já mostramos que os pares de fórmulas da proposição abaixo satisfazem a relação de consequência lógica probabilística. A seguir mostramos que elas não satisfazem a relação de consequência lógica informacional.

Proposição 6.4.8: Seja $\Gamma = \{\varphi_1, ..., \varphi_n\}$ o conjunto de premissas de cada item abaixo e seja φ_m a sua conclusão, para n, $m \in \mathbb{N}$. Então:

a: $\varphi \wedge \psi \nVdash \varphi$.

b: $\varphi \rightarrow \psi, \varphi \nVdash \psi$.

c: $\varphi \rightarrow \psi, \neg\psi \nVdash \neg\varphi$.

d: $\varphi \rightarrow \psi, \psi \rightarrow \gamma \nVdash \varphi \rightarrow \gamma$.

e: $\varphi \nVdash \psi \rightarrow \varphi$.

f: $\psi \nVdash \varphi \vee \psi$.

g: $\gamma \vee \varphi, \neg\gamma \vee \psi \nVdash \varphi \vee \psi$.

Demonstração

Pode-se mostrar, para cada item, que existe f tal que $I(f)(\Gamma) = 0$ e $I(f)(\varphi_m) > 0$. ■

A proposição acima mostra que as regras de inferência de grande parte dos sistemas formais lógicos não são válidas informacionalmente. Em especial, os dois últimos itens ilustram que a regra da expansão e a regra do corte, tal como na Definição 2.2.12, duas das regras de inferência de qualquer sistema formal lógico de primeira ordem clássico, não possuem validade geral em termos informativos.

O único caso que pode ocorrer da informação das premissas ser maior do que a informação na conclusão é quando o conjunto de premissas possui valor probabilidade zero. Este foi exatamente o caso utilizado na demonstração da proposição acima. Considerando-se a perspectiva informacional e o arcabouço conceitual que estamos considerando, pode-se justificar este resultado para estes casos.

Para tal justificativa, consideremos a Proposição 6.4.8f. Quando a probabilidade da premissa for nula, faz sentido dizer que $\varphi \vee \psi$ não é consequência lógica informacional de ψ. Enquanto premissa nada informa, nada diz sobre o universo, a conclusão informa algo a respeito dele. A sentença “caiu o número sete” no jogo de dado, por exemplo, nada diz sobre o que ocorreu no mundo. No entanto, a sentença “caiu número sete ou caiu par” possui quantidade de informação, uma vez que reduz incerteza: pode ter caído número dois, quatro ou seis. Assim, tenho na conclusão algo que não tinha na premissa. Ou seja, houve um ganho informacional, unicamente porque a quantidade de informação da premissa é nula.

Da proposição acima concluímos ainda que argumentos como os do Exemplo 6.4.3 são inválidos na perspectiva informacional da consequência lógica. No caso do Exemplo 6.4.3a, pode-se dizer que ele possui a forma de *modus ponens* que, como mostramos na Proposição 6.4.8b, é inválido. Para este argumento possuir a forma de *modus ponens*, é necessária uma adequação das suas premissas: podemos entender que a sentença “Todo homem é mortal” é equivalente a “Se algo é homem, então algo é mortal” que, por sua vez, é equivalente a “Algo não é homem ou é mortal”. Quando os dois predicados desta sentença forem mutuamente exclusivos, tem-se que “Algo não é homem ou algo não é homem”, que é equivalente a “Algo não é homem”. Quando o termo “algo” for aplicado a Sócrates, como no Exemplo 6.4.3a, as premissas expressam que Sócrates não é homem e

Sócrates é homem. Assim, a conclusão do argumento, ao afirmar que Sócrates é mortal, pode possuir quantidade de informação maior do que a quantidade de informação do conjunto de premissas, que é nulo neste caso. Logo, a conclusão não é consequência lógica informacional do conjunto de premissas.

O próximo resultado explicita quais são os argumentos válidos segundo a perspectiva informacional da consequência lógica.

Teorema 6.4.9

Seja $\Gamma = \{\varphi_1, ..., \varphi_n\}$. Então $\Gamma \Vdash \varphi \Leftrightarrow I(\varphi) = 0$, para toda f, ou $(\varphi_1 \wedge ... \wedge \varphi_n) \equiv \varphi$.

Demonstração

(⇒): seja $\Gamma \Vdash \varphi$. Suponha que $I(\varphi) \neq 0$, para alguma f e $(\varphi_1 \wedge ... \wedge \varphi_n) \not\equiv \varphi$. Então pode-se mostrar a existência de f' tal que $I(f')(\Gamma) < I(f')(\varphi)$ contrariando a hipótese inicial: $I(f')(\Gamma) = 0$ e $I(f')(\varphi) = I(f)(\varphi)$. Assim, quando $\Gamma \Vdash \varphi$ tem-se que: $I(\varphi) \neq 0$, para alguma $f \Rightarrow (\varphi_1 \wedge ... \wedge \varphi_n) \equiv \varphi$, ou seja: $I(\varphi) = 0$, para toda f, ou $(\varphi_1 \wedge ... \wedge \varphi_n) \equiv \varphi$.

(⇐): Caso 1: $I(\varphi) = 0$, para toda $f \Rightarrow I(\Gamma) \geq I(\varphi)$, para toda f e todo $\Gamma \Rightarrow_{Def6.4.5}\Rightarrow \Gamma \Vdash \varphi$.

Caso 2: $(\varphi_1 \wedge ... \wedge \varphi_n) \equiv \varphi \Leftrightarrow_{Def5.4.8}\Leftrightarrow P(\varphi_1 \wedge ... \wedge \varphi_n) = P(\varphi)$, para toda $f \Leftrightarrow_{Def6.3.1}\Leftrightarrow P(\Gamma) = P(\varphi)$, para toda $f \Rightarrow_{Teo6.2.4.c}\Rightarrow I(\Gamma) = I(\varphi)$, para toda $f \Rightarrow_{Def6.4.5}\Rightarrow \Gamma \Vdash \varphi$. ■

No resultado a seguir explicitamos que a Lógica subjacente à consequência lógica informacional não é um sistema formal clássico.

Corolário 6.4.10

a: $I(\varphi) > 0$, para alguma $f \Rightarrow \bot_{\mathbb{P}} \nVdash \varphi$.

b: $\Gamma \Vdash \top_{\mathbb{P}}$.

c: $\Gamma \Vdash \bot_{\mathbb{P}}$.

d: $I(\varphi) > 0$, para alguma $f \Rightarrow \top_{\mathbb{P}} \nVdash \varphi$.

Demonstração

a: Seja $P(\varphi) = ½$. Então $I(\bot_{\mathbb{P}}) = 0 < I(\varphi) = 1$.

b: Como $I(\top_{\mathbb{P}}) = 0$, tem-se que $I(\Gamma) \geq I(\top_{\mathbb{P}})$.

c: Como $I(\bot_{\mathbb{P}}) = 0$, tem-se que $I(\Gamma) \geq I(\bot_{\mathbb{P}})$.

d: Seja $P(\varphi) = ½$. Então $I(\top_{\mathbb{P}}) = 0 < I(\varphi) = 1$. ■

Este resultado pode ser mostrado a partir do teorema anterior. Seu primeiro item explicita que de uma fórmula probabilisticamente contraditória não se segue informacionalmente qualquer fórmula. De fato, apenas fórmulas com informação zero são consequência lógica de uma fórmula contraditória. Nenhuma fórmula contingente, ou seja, informativa, é consequência lógica informacional de uma fórmula contraditória.

Do item em questão concluímos ainda que a *Lógica Clássica não é a Lógica subjacente à consequência lógica informacional*. Isso porque, nos sistemas formais lógicos clássicos, uma contradição gera qualquer fórmula. Considerando que as Lógicas Complementares, como a Lógica Modal, preservam os mesmos princípios da Lógica Clássica, também concluímos que *Nenhuma Lógica Complementar é a Lógica subjacente à consequência lógica informacional*. Restam como candidatas a essa Lógica as Lógicas Heterodoxas, como, por exemplo, os sistemas Intuicionistas ou Paraconsistentes.

Conforme visto na Seção 1.5, nas Lógicas Intuicionistas a negação possui algumas características próprias. Tais particularidades não permitem, por exemplo, o recurso às provas por Redução ao Absurdo, dado que fórmulas como "$\varphi \leftrightarrow \neg\neg\varphi$" não são válidas nestes sistemas. Entretanto, temos que $\neg\neg\varphi \Vdash \varphi$ e $\varphi \Vdash \neg\neg\varphi$ e, pelo Teorema 6.4.11b abaixo, temos que $\Vdash \varphi \leftrightarrow \neg\neg\varphi$. Assim, a *Lógica Intuicionista não pode ser subjacente à consequência lógica informacional*.

Acreditamos intuitivamente que a Lógica subjacente à consequência lógica informacional é, no mínimo, Paraconsistente, *lato-sensu*. Isso porque, nessa noção de Consequência não vale, por exemplo, o *Princípio da Explosão*, ou seja, de uma contradição não se segue qualquer coisa.

O Corolário 6.4.10a reforça a concordância da consequência lógica informacional com a noção intuitiva de consequência informacional. Ao dizer que "hoje é terça e hoje não é terça" não posso concluir disso nem que hoje é terça nem que hoje não é terça. Tampouco posso concluir nada que possua outro conteúdo, por exemplo, que "a bengala está no canto". Nas premissas eu nada digo a respeito de bengalas ou qualquer outra coisa que não seja hoje é terça.

O segundo item é compartilhado pela consequência lógica probabilística e pela veritativo-funcional, mas por motivos distintos. Informacionalmente, uma fórmula probabilisticamente válida é consequência de qualquer fórmula porque o seu valor é mínimo. Assim, ela nada pode informar a mais do que qualquer outra fórmula. Em termos probabilísticos e veritativo-funcionais, ela é consequência lógica de qualquer fórmula porque o seu valor é máximo.

O terceiro item, como o primeiro, destoa das outras duas noções de consequência lógica. Nestas, uma fórmula probabilisticamente válida é consequência unicamente de outra fórmula probabilisticamente válida. Aqui, ela é consequência de qualquer fórmula. Isso se deve ao fato de que a noção informacional não distingue as fórmulas válidas das contraditórias.

O quarto resultado expressa uma diferença entre a consequência lógica informacional e as outras noções de consequência, a ser mostrada no Teorema 6.4.11f abaixo.

Teorema 6.4.11

a: $\varphi \Vdash \psi$ e $\psi \Vdash \gamma \Rightarrow \varphi \Vdash \gamma$.

b: $\varphi \equiv \psi \Rightarrow\ \Vdash \varphi \leftrightarrow \psi$.

c: $\Gamma, \varphi \Vdash \psi \not\Leftrightarrow \Gamma \Vdash \varphi \rightarrow \psi$.

d: $\Vdash \varphi$ e $\Vdash \psi \not\Leftrightarrow \varphi \equiv \psi$.

e: $\Gamma \Vdash \varphi \not\Leftrightarrow f(\Gamma) \subseteq f(\varphi)$, para toda f.

f: $\Gamma \Vdash \varphi \not\Leftrightarrow \Gamma \vDash_{\mathcal{V}} \varphi$.

Demonstração

a: $\varphi \Vdash \psi$ e $\psi \Vdash \gamma \Rightarrow_{\text{Def6.4.5}}\Rightarrow \mathrm{I}(\varphi) \geq \mathrm{I}(\psi) \geq \mathrm{I}(\gamma)$, para toda $f \Rightarrow \mathrm{I}(\varphi) \geq \mathrm{I}(\gamma)$, para toda f $\Rightarrow_{\text{Def6.4.5}}\Rightarrow \varphi \Vdash \gamma$.

b: $\varphi \equiv \psi \Rightarrow_{\text{Teo5.4.9b}}\Rightarrow \varphi \leftrightarrow \psi$ é $\top_{\mathbb{P}} \Rightarrow_{\text{Teo6.3.7a}}\Rightarrow \vDash_{\mathbb{P}} \varphi \leftrightarrow \psi \Rightarrow_{\text{Teo6.4.7d}}\Rightarrow\ \Vdash \varphi \leftrightarrow \psi$.

c: **($\not\Rightarrow$):** seja $\Gamma = \varnothing$. $\bot_{\mathbb{P}} \not\Vdash \psi$, mas $\Vdash \bot_{\mathbb{P}} \rightarrow \psi$.

($\not\Leftarrow$): $\not\Vdash \psi \rightarrow \bot_{\mathbb{P}}$, mas $\psi \Vdash \bot_{\mathbb{P}}$.

d: **($\not\Rightarrow$):** $\Vdash \bot_{\mathbb{P}}$ e $\Vdash \top_{\mathbb{P}}$, mas $\bot_{\mathbb{P}} \not\equiv \top_{\mathbb{P}}$.

($\not\Leftarrow$): $\varphi \rightarrow \psi \equiv \neg\varphi \vee \psi$, mas não ocorre que: $\vDash_{\mathbb{P}} \varphi \rightarrow \psi$ e $\vDash_{\mathbb{P}} \neg\varphi \vee \psi$.

e: **($\not\Rightarrow$):** $\top_{\mathbb{P}} \Vdash \bot_{\mathbb{P}}$, mas $f(\top_{\mathbb{P}}) = \mathrm{U} \not\subseteq f(\bot_{\mathbb{P}}) = \varnothing$.

($\Leftarrow\!\!\!\!/\,$): Seja $f(\varphi) \neq \emptyset$. Então $f(\bot_{\mathbb{P}}) \subseteq f(\varphi)$, mas $\bot_{\mathbb{P}} \not\Vdash \varphi$.

f: **($\not\Rightarrow$):** $\varphi \vee \neg\varphi \Vdash \varphi \wedge \neg\varphi$, mas $\varphi \vee \neg\varphi \not\vDash_{\mathcal{V}} \varphi \wedge \neg\varphi$.

($\not\Leftarrow$): $\psi \vDash_{\mathcal{V}} \varphi \vee \psi$, mas $\psi \not\Vdash \varphi \vee \psi$. ■

Este teorema expressa algumas características próprias da consequência lógica informacional. Grande parte das equivalências acima foi mostrada para as versões veritativo-funcional e probabilística. A recíproca do Teorema 6.4.11b, por exemplo, não pode ser mostrada para a versão informacional: $\Vdash \varphi \leftrightarrow \psi \not\Rightarrow \varphi \equiv \psi$. O seguinte caso serve de exemplo: $\Vdash \bot_{\mathbb{P}} \leftrightarrow \top_{\mathbb{P}}$, mas $\bot_{\mathbb{P}} \not\equiv \top_{\mathbb{P}}$.

O Teorema 6.4.11c é o correspondente semântico do Teorema da Dedução e o Teorema 6.4.11f expressa que a relação de consequência lógica informacional e a relação de consequência lógica veritativo-funcional são distintos.

O último item mostra que, nas situações em que um conjunto de premissas é informativo, o evento associado a elas está contido no evento associado à conclusão. Nestas situações, em todos os itens da Proposição 6.4.8, a quantidade de informação da conclusão sempre seria menor do que a quantidade de informação do conjunto de premissas. Isto significaria que o conjunto de premissas seria mais preciso, arriscado, menos provável, portanto, mais informativo do que a conclusão.

Mostramos a seguir que a consequência lógica informacional não é uma consequência lógica Tarskiana.

Teorema 6.4.12

a: $\varphi \in \Gamma \not\Rightarrow \Gamma \Vdash \varphi$.

b: $\Gamma \subseteq \Delta$ e $\Gamma \Vdash \varphi \not\Rightarrow \Delta \Vdash \varphi$.

c: $\Delta \Vdash \psi$, para cada $\psi \in \Gamma$, e $\Gamma \Vdash \varphi \Rightarrow \Delta \Vdash \varphi$.

Demonstração: seja *f* uma situação qualquer.

a: Sejam $P(\Gamma) = 0$ e $P(\varphi) = ½$. Então $I(\Gamma) < I(\varphi)$.

b: Sejam $\Gamma = \{\varphi\}$, $\Delta = \{\varphi, \neg\varphi\}$ e $I(\varphi) > 0$. Então $\Gamma \Vdash \varphi$ e $\Delta \not\Vdash \varphi$.

c: Caso 1: $I(\Delta) = 0 \Rightarrow_{Hip}\Rightarrow I(\psi) = 0$, para cada $\psi \in \Gamma \Rightarrow I(\Gamma) = 0 \Rightarrow_{Hip}\Rightarrow I(\varphi) = 0 \Rightarrow I(\Delta) \geq I(\varphi)$.

Caso 2: $I(\Delta) \neq 0 \Rightarrow_{Teo6.2.4b;Hip}\Rightarrow P(\Delta) \leq P(\psi)$, para cada $\psi \in \Gamma$. Neste caso, pode-se mostrar que $f'(\Delta) \subseteq f'(\Gamma) \subseteq f'(\varphi)$, para toda $f' \Rightarrow_{Teo6.3.7c;Def6.3.4}\Rightarrow P(\Delta) \leq P(\Gamma) \leq P(\varphi)$ $\Rightarrow_{Teo6.2.4a}\Rightarrow I(\Delta) \geq I(\varphi)$. ■

De acordo com o teorema acima, a consequência lógica informacional não é reflexiva e tampouco monotônica, embora satisfaça a propriedade da transitividade.

Em suma, neste capítulo desenvolvemos uma análise informacional da consequência lógica. Introduzimos uma medida informacional para as fórmulas de *LSC* e para conjuntos de fórmulas, a partir da qual definimos a consequência lógica informacional, depois de expor uma definição probabilística da consequência lógica. Concluímos que, além de ser Tarskiana, a consequência lógica probabilística é satisfeita pelos mesmos elementos que satisfazem a consequência lógica veritativo-funcional. Mostramos também que as fórmulas informacionalmente vazias são consequência lógica informacional do conjunto vazio e que as fórmulas que são consequência lógica informacional de um dado conjunto de fórmulas cuja quantidade de informação é sempre nula são consequência lógica informacional do conjunto vazio. Além disso, algumas regras de inferência dos sistemas formais lógicos clássicos não possuem validade geral na perspectiva informacional da consequência lógica. Isto nos levou a mostrar que a consequência lógica informacional é distinta da consequência lógica probabilística. Provamos que uma fórmula é consequência lógica informacional de um dado conjunto de fórmulas sse o dado conjunto de premissa é probabilisticamente equivalente à conclusão ou a conclusão é informacionalmente nula. Por fim, mostramos que a consequência lógica informacional não é uma consequência lógica Tarskiana e que a Lógica subjacente a ela é, no mínimo, Paraconsistente, *lato-sensu*.

Considerações Finais

Neste trabalho buscamos desenvolver uma análise da consequência lógica a partir de uma perspectiva quantitativa da informação. Na primeira parte, que consiste dos três capítulos iniciais, fazemos um estudo crítico de algumas das principais concepções usuais de consequência lógica. Na segunda parte, constituída dos três últimos capítulos, desenvolvemos a nossa proposta. Ressaltamos, a seguir, os principais resultados obtidos no decorrer do trabalho.

(CF1) As três características elementares da consequência lógica são a necessidade, a formalidade e a anterioridade (Seção 1.2).

A *necessidade* indica o caráter de confiabilidade, de garantia absoluta da conclusão a partir das premissas. Para que uma sentença seja consequência lógica de um dado conjunto de sentenças, a relação ou a inferência deve ser mantida em qualquer circunstância logicamente possível.

A *formalidade* explicita a ideia de que a consequência lógica é identificada ou definida a partir da estrutura ou da forma dos seus elementos constituintes e não a partir do seu conteúdo.

A *anterioridade* ou *aprioridade* pressupõe a independência da consequência lógica de conhecimentos individuais ou prévios, do grau informacional sobre o mundo ou significado dos termos, de relações de conteúdo ou algum tipo de experiência referente aos fatos expressos pelas premissas e conclusão.

(CF2) A consequência lógica sintática não satisfaz a característica da necessidade e da anterioridade (Capítulo 2).

(CF2.1) A existência de sistemas formais lógicos diferentes com relações de consequência logica distintas desconsidera a necessidade e a anterioridade (Seção 2.3).

Sistemas formais lógicos diferentes podem possuir relação de consequência lógica distintos, ainda que a noção de consequência lógica subjacente seja a mesma. Ou seja, uma fórmula pode ser consequência lógica de um dado conjunto de fórmulas em um sistema formal lógico mas não o ser em outro. No entanto, de acordo com a característica da necessidade, a consequência lógica entre uma sentença e um dado conjunto de sentenças deve ser mantida em qualquer circunstância, em qualquer sistema formal lógico. Além disso, a pressuposição do conhecimento do sistema formal lógico para a definição de uma consequência lógica desconsidera a característica da anterioridade.

(CF2.2) A existência de diferentes noções de consequência lógica igualmente apropriadas desconsidera a característica da necessidade e da anterioridade (Seção 2.4).

Podem existir diferentes noções de consequência lógica, com propriedades distintas, monotônicas ou não-monotônicas, por exemplo. Disto resulta a possibilidade de que uma fórmula seja consequência lógica de um dado conjunto de fórmulas segundo uma concepção, mas não segundo outra, o que desconsideraria a característica da necessidade. Além disso, a pressuposição do conhecimento das propriedades de cada consequência lógica para saber se uma fórmula é consequência lógica de um dado conjunto de fórmulas desconsidera a característica da anterioridade.

(CF3) A consequência lógica semântica, em suas variadas versões, não satisfaz alguma das características da consequência lógica (Capítulo 2).

(CF3.1) A abordagem modal desconsidera a característica da formalidade (Exemplo 3.2.2a) e da anterioridade (Exemplo 3.2.2b).

A validade de argumentos como o do Exemplo 3.3.2a é garantida graças ao conteúdo das suas sentenças, uma condição extra-lógica. Argumentos com a mesma estrutura podem apresentar premissa verdadeira e conclusão falsa, sendo intuitivamente considerados inválidos, como é o caso do Exemplo 3.2.3.

Por outro lado, a validade do argumento do Exemplo 3.2.2b depende, além do conteúdo das sentenças, de um conhecimento prévio, qual seja, a identificação do significado ou da referência dos termos envolvidos, o que nos faz concluir que a abordagem modal não satisfaz a anterioridade.

(CF3.2) A abordagem formal, em sua versão substitucional, desconsidera as características de necessidade e de formalidade (Exemplos 3.3.2 a 3.3.4).

A necessidade e a formalidade não são satisfeitas por esta abordagem formal devido ao problema dos termos lógicos. A validade de argumentos como o do Exemplo 3.3.2, em uma dada linguagem, depende da escolha destes termos.

O Exemplo 3.3.4 mostra que a validade de um argumento na abordagem formal também pode depender da variação da complexidade da linguagem subjacente. Isso a faz desconsiderar a característica da necessidade. O mesmo exemplo serve para ilustrar que tal abordagem desconsidera a característica da formalidade, uma vez que a validade do argumento não depende da distribuição dos termos, mas de um aspecto geográfico, uma relação exta-lógica.

(CF3.3) A abordagem formal, em sua versão interpretacional, desconsidera as características da necessidade e da anterioridade (Exemplos 3.3.7 a 3.3.9).

O Exemplo 3.3.7 mostra que, nesta versão da abordagem formal, a validade de certos argumentos depende da escolha dos temos lógicos, o que a faz desconsiderar a característica da necessidade. Já o Exemplo 3.3.8 mostra que a validade de um argumento pode depender da análise do mundo *real*, desconsiderando a anterioridade.

Por sua vez, o Exemplo 3.3.9 mostra que, mesmo adotando certos símbolos como essencialmente lógicos, a validade do argumento exemplificado depende de um fator extra-lógico, qual seja, o tamanho do universo, o que faz com que esta versão da abordagem formal desconsidere as características da necessidade e da anterioridade.

(CF3.4) A abordagem modal/formal desconsidera as características da necessidade e da anterioridade (Exemplos 3.4.11 e 3.4.12).

Os exemplos em questão ilustram que a validade de certos argumentos depende de fatores extra-lógicos, como a garantia da infinitude do universo, o que a faz desconsiderar as características da necessidade e da anterioridade. Tal perspectiva é construída com base na teoria de conjuntos do tipo ZF, nas quais o axioma da infinitude garante que o universo contém um número infinito de objetos.

(CF4) O valor probabilidade de uma implicação material, $\varphi \rightarrow \psi$, em geral, não equivale ao valor da implicação probabilística, $\psi|\varphi$ (Teorema 5.3.4f).

Este resultado acima é derivado de nossa distinção conceitual entre a implicação probabilística e a implicação material probabilística.

(CF5) A implicação material não captura a noção intuitiva de implicação quando esta é entendida como a relação de causalidade clássica (Exemplo 5.3.5).

Embora, no referido exemplo, utilizamos a implicação material probabilística como referência, os mesmos resultados podem ser obtidos para a implicação material veritativo-funcional. No mesmo exemplo mostramos que a implicação probabilística parece capturar a noção de causalidade clássica.

Nosso conceito de implicação probabilística parece capturar a noção de causalidade clássica. Algumas consequências desta nossa abordagem parecem relevantes.

Lembramos que o condicional probabilístico, φ implica probabilisticamente ψ, é denotado em $L^{|}$ por "$\psi|\varphi$" e $P(\psi|\varphi) = p(f(\psi)|f(\varphi))$. Além disso, as definições de $S_{\mathbb{P}}$, por ora aplicadas a $L^{|}$, foram expandidas a fim de considerar também o conectivo da implicação probabilística.

Uma das dificuldades de $S_{\mathbb{P}}$ para $L^{|}$ está na atribuição dos eventos correspondentes às fórmulas implicativas probabilísticas em uma situação. Sem essa atribuição, não há como determinar o valor probabilidade de certas fórmulas implicativas probabilísticas em $S_{\mathbb{P}}$, como por exemplo, da fórmula $(\varphi|\psi)|\varphi$ quando $0 < P(\varphi), P(\psi) < 1$. Isso porque não definimos, em $S_{\mathbb{P}}$, o evento correspondente a $\varphi|\psi$ em uma dada situação. Tal definição necessitaria de algum recurso conceitual, sendo que, para os propósitos deste trabalho, é desnecessário fazê-lo.

Apesar da dificuldade apontada acima, é possível obter alguns resultados surpreendentes a respeito de certas fórmulas implicativas probabilísticas em determinadas situações. Esse é o caso dos seguintes exemplos:

a: $(\varphi|\psi)|\varphi$ não é $\top_{\mathbb{P}}$

Seja f tal que $p(f(\varphi)) = 0$. Assim, $P((\varphi|\psi)|\varphi) = p(f(\varphi|\psi)|f(\varphi)) = 0$.

b: $\psi|\varphi \vee \varphi|\psi$ não é $\top_{\mathbb{P}}$

Seja f tal que $p(f(\varphi)) = p(f(\psi)) = 0$. Assim, $P(\psi|\varphi) = P(\varphi|\psi) = 0$.

c: $(\psi|\varphi)|\neg\varphi$ não é $\top_{\mathbb{P}}$

Seja f tal que $p(f(\varphi)) = 1$. Assim, $P((\psi|\varphi)|\varphi) = p(f(\psi|\varphi)|f(\varphi)) = 0$.

d: $\varphi|\varphi$ não é $\top_{\mathbb{P}}$

Seja f tal que $p(f(\varphi)) = 0$. Assim, $P(\varphi|\varphi) = p(f(\varphi)|f(\varphi)) = 0$.

As quatro fórmulas implicativas probabilísticas acima correspondem às seguintes fórmulas implicativas materiais:

a' $\varphi \to (\psi \to \varphi)$

b' $(\varphi \to \psi) \vee (\psi \to \varphi)$

c' $\neg\varphi \to (\varphi \to \psi)$

d' $\varphi \to \varphi$.

As três primeiras fórmulas implicativas probabilísticas acima expressam que os paradoxos da implicação material não são aplicados à implicação probabilística. O terceiro resultado é muito expressivo. De algum modo, ele indica que, de uma contradição não se segue qualquer coisa. Abstrairíamos daí que, em termos da implicação probabilística, a inconsistência de um sistema formal não leva necessariamente à sua trivialidade. Sistemas que adotam esta noção de implicação são algum tipo de sistemas paraconstistentes.

A quarta fórmula indica outra peculiaridade da implicação probabilística: nem sempre uma fórmula implica probabilisticamente a si mesma. O único caso em que uma fórmula do tipo $\varphi|\varphi$ é válida é aquele em que φ também é válida. No entanto, quando φ não é válida, então $P(\varphi|\varphi) \neq 1$ unicamente nas situações em que $P(\varphi) = 0$. Assim, o que pareceria inicialmente uma desvantagem da implicação probabilística, pode ser utilizado em seu favor: quando a probabilidade de ocorrência de uma fórmula é nula, também é apropriado dizer que a probabilidade de ocorrência dessa fórmula ocorrer, dada a ocorrência dela mesma, é zero.

A seguir comentamos outros resultados interessantes da implicação probabilística.

e: $\perp_{\mathbb{P}}|\varphi$ é $\perp_{\mathbb{P}}$;

f: $\varphi|\perp_{\mathbb{P}}$ é $\perp_{\mathbb{P}}$;

g: $\top_{\mathbb{P}}|\varphi$ é $\top_{\mathbb{P}}$;

h: $P(\varphi|\top_{\mathbb{P}}) = P(\varphi)$.

Estes resultados são consequência imediata da definição de implicação probabilística e da definição de probabilidade condicional.

Os dois primeiros itens, que parecem intuitivamente adequados, mostram que, segundo a concepção de implicação probabilística, uma contradição jamais implica ou é implicada por uma fórmula. Dado que $P(\bot_{\mathbb{P}}) = 0$, para toda situação, pode-se interpretar "$\bot_{\mathbb{P}}$" como a indicação do impossível, significando a inexistência de acontecimentos, pois $f(\bot_{\mathbb{P}}) = \varnothing$. Um dos modos de expressar as fórmulas probabilísticas inconsistentes é através da contradição, indicando a ocorrência de um evento e a não ocorrência desse mesmo evento, ou seja, do seu complemento. Essa conjunção, nos termos de $\mathbb{P}$, resulta no vazio.

A primeira fórmula diz que φ implica probabilisticamente o impossível. Isso quer dizer: a ocorrência de um dado evento implica probabilisticamente a ocorrência de nenhum evento, ou de outro modo: a ocorrência de nenhum evento, dada a ocorrência de algum evento. Isso é impossível, pois, se algo ocorre, então a probabilidade de nada ocorrer é nula. Caso exótico seria aquele em que $P(\varphi) = 0$. Mas, também nesse caso, $P(\bot_{\mathbb{P}}|\varphi) = 0$. Poderíamos ainda interpretar a fórmula em questão como: a ocorrência de um dado evento implica probabilisticamente a ocorrência e a não ocorrência de algo, ou a probabilidade de algo acontecer e não acontecer, dada a ocorrência de um dado evento.

Já a segunda fórmula significa o contrário da primeira: a não-ocorrência de eventos implica probabilisticamente a ocorrência de um dado evento, ou a probabilidade de algo ter ocorrido, dado que nada ocorreu é realmente impossível. Esta fórmula garante que, em termos de implicação probabilística, de uma contradição não se segue qualquer coisa. A rigor, de uma contradição nada é implicado probabilisticamente.

O terceiro e quarto resultados parecem intuitivamente adequados, como procuramos mostrar a seguir. Dado que $P(\top_{\mathbb{P}}) = 1$, para toda situação, pode-se interpretar "$\top_{\mathbb{P}}$" como a indicação do necessário, significando a existência de qualquer acontecimento. Um dos modos de expressar as fórmulas probabilisticamente válidas é através do terceiro excluído, indicando a ocorrência de um evento ou a não ocorrência desse mesmo evento, ou seja, do seu complemento. Essa disjunção, nos termos de $\mathbb{P}$, resulta no próprio espaço amostral, ou seja, no universo de acontecimentos.

A terceira fórmula significa que φ implica probabilisticamente o necessário. Isso quer dizer que a ocorrência de um dado evento implica probabilisticamente a ocorrência de algum evento, o que é óbvio. Ou pode significar a ocorrência de algum evento, dada a ocorrência de determinado evento. Como é necessário, o valor probabilidade da fórmula deveria ser sempre máximo, como indica o resultado acima.

A quarta fórmula expressa que a ocorrência de algo implica probabilisticamente a ocorrência de um dado evento. Ou a probabilidade de ocorrer um dado evento, dado que ocorreu algum evento. Intuitivamente, é óbvio que a probabilidade dessa fórmula depende da probabilidade de φ, como definido pelo resultado em questão.

Outros resultados próprios da implicação probabilística são os seguintes, mostrados com base no Exemplo 5.3.3:

i: $\psi|\varphi \not\equiv \neg\varphi \vee \psi$

$P(f(\Sigma_1))(A_3|A_1) = 1/3$ e $P(f(\Sigma_1))(\neg A_1 \vee A_3) = 2/3$.

j: $\psi|\varphi \not\equiv \neg(\varphi \wedge \neg\psi)$

$P(f(\Sigma_1))(A_3|A_1) = 1/3$ e $P(f(\Sigma_1))(\neg(A_1 \wedge \neg A_3)) = 2/3$.

k: $\psi|\varphi \not\equiv \neg\varphi|\neg\psi$

$P(f(\Sigma_1))(A_3|A_1) = 1/3$ e $P(f(\Sigma_1))(\neg A_1|\neg A_3) = 0$.

l: $\psi|\varphi \not\equiv \varphi|\psi$

$P(f(\Sigma_1))(A_3|A_1) = 1/3$ e $P(f(\Sigma_1))(A_3|A_1) = ¼$.

Os dois primeiros resultados acima são consequência imediata das definições do valor probabilidade da implicação material probabilística e implicação probabilística, dado que $\psi|\varphi \not\equiv \varphi \rightarrow \psi \equiv \neg\varphi \vee \psi \equiv \neg(\neg\varphi \wedge \neg\psi)$. O terceiro resultado mostra que a contraposição não é válida para o condicional probabilístico. Já o último resultado é esperado, pois o fato de uma fórmula implicar uma segunda não significa que esta também implique aquela.

O primeiro resultado acima mostra ainda que a implicação probabilística é capaz de capturar a noção de quantidade de informação da TMC nesta situação particular. Ao contrário da implicação material probabilística, como discutido no Teorema 6.2.5a, nem sempre a quantidade de informação nas duas sentenças coincide. Considerando a situação do jogo de um dado não viciado, a quantidade de informação na sentença “se cai primo, então cai par”, por exemplo, não é, intuitivamente falando, igual à da sentença “não cai primo ou cai par”. A quantidade de informação da sentença implicativa, nesta situação, é bem maior do que a informação presente na sentença disjuntiva. É muito difícil, ou muito pouco provável, a ocorrência de par, na hipótese da ocorrência de primo; porém, é bem mais provável que ocorra o evento associado à sentença “não cai primo ou cair par”. Nesta situação, se associarmos a sentença “cai primo” à fórmula A_1 e a sentença “cai par” a A_2, então $P(A_1|A_2) \neq P(\neg A_2 \vee A_2)$.

No entanto, no comentário sobre o Teorema 6.2.5a, apresentamos o exemplo do extraterrestre: na sentença: “se há vida extraterrestre, então eu sou trípide”, o falante não está querendo dizer “não há vida extraterrestre ou eu sou um trípide”. A sentença implicativa, neste contexto, significa categoricamente que “não há vida extraterrestre”. Assim, considerando que sentenças com o mesmo conteúdo devem possuir probabilidades iguais, neste caso, deveríamos ter que $P(A_1|A_2) \neq P(\neg A_2)$. Mas $P(A_1|A_2) = 0$, dado que $P(A_1) = 0$. Se, na implicação acima, o conteúdo informacional consiste em dizer que não há vida extraterrestre, então $P(\neg A_2) = 1$. Isso contraria o que desejávamos.

Em suma, apesar de o condicional material possuir algumas vantagens sobre o condicional probabilístico, especialmente quando tratamos da noção intuitiva de implicação, ele também apresenta algumas características que podem ser consideradas muito desvantajosas.

(CF6) O conjunto de fórmulas válidas segundo $S_{\mathbb{P}}$ e segundo $S_{\mathcal{V}}$ é exatamente o mesmo (Teorema 5.4.6b)

Deste resultado obtém-se que os paradoxos da implicação material são mantidos na perspectiva probabilística, conforme mostrado no comentário feito após o Teorema 5.4.6. Disto resulta, dentre outras coisas, que nem sempre há relação entre a razoabilidade de uma fórmula e a relação de conteúdo entre seus constituintes, conforme mostrado no Teorema 5.3.7, tomando como referência a implicação material probabilística.

Desta equivalência entre as duas semânticas também podemos provar o Meta-Teorema da Completude Fraco de *LSC* relativa a $S_{\mathbb{P}}$: o conjunto de teoremas de *LSC* é igual ao conjunto de fórmulas probabilisticamente válidas. Disso resulta que $S_{\mathbb{P}}$ é uma semântica para *LSC* (Definição 5.4.4a).

(CF7) $S_{\mathbb{P}}$ não é uma semântica clássica para *LSC* (Definição 5.4.4b;Teorema 5.4.3).

$S_{\mathbb{P}}$ apresenta algumas características próprias que a distinguem de $S_{\mathcal{V}}$. Em $S_{\mathbb{P}}$ não vale, por exemplo, o princípio da bivalência. Alguns desdobramentos dos princípios do terceiro excluído e da não-contradição tampouco são satisfeitos por $S_{\mathbb{P}}$, como mostrado no comentário sobre o Teorema 5.4.3. Além disso, em $S_{\mathbb{P}}$ não valem alguns resultados possíveis de serem provados em $S_{\mathcal{V}}$, como mostrado no Teorema 5.3.6 e no Teorema 5.4.9.

(CF8) $S_{\mathbb{P}}$ não captura o conteúdo informacional das sentenças (Teorema 6.2.4c).

No Teorema 6.2.4c mostramos que fórmulas com o mesmo valor probabilidade possuem o mesmo valor informacional. Nosso objetivo, neste trabalho, não é fazer uma análise do conteúdo informacional de uma sentença. Não temos a intenção de tratar do aspecto qualitativo da informação, ou do problema semântico, conforme exposto na Seção 4.2. No entanto, para mostrar que a perspectiva informacional em questão trata apenas do problema técnico, sendo incapaz de capturar o problema semântico, justificamos a seguir dois *postulados da informação semântica*, comparados à probabilidade de uma sentença. Explicitaremos que a perspectiva informacional em questão não satisfaz estes postulados. A expressão "$I_S(f(\Sigma))(\varphi)$" significa "o conteúdo informacional da fórmula φ de *L* na situação $f(\Sigma)$".

(P1): $I_S(f(\Sigma))(\varphi)$ é a mesma $I_S(f(\Sigma))(\psi) \Rightarrow P(f(\Sigma))(\varphi) = P(f(\Sigma))(\psi)$.
(P2): $P(f(\Sigma))(\varphi) = P(f(\Sigma))(\psi) \nRightarrow I_S(f(\Sigma))(\varphi)$ é a mesma $I_S(f(\Sigma))(\psi)$.

O postulado (P1) determina que fórmulas com o mesmo conteúdo informacional devem possuir igual valor probabilidade. Se duas fórmulas dizem a mesma coisa, então parece evidente que a chance de ocorrência do que cada uma diz deve ser igual. Assim, se as sentenças "João ama Maria" e "Maria é amada por João" possuírem o mesmo conteúdo informacional, ou seja, se disserem o mesmo, então a probabilidade de ambas deve ser igual.

Quando $I_S(\varphi)$ não é a mesma $I_S(\psi)$, o valor probabilidade de φ e ψ pode ser tanto igual quanto diferente em uma dada situação. No exemplo do lance de um dado, os eventos "par" e "ímpar" possuem conteúdos distintos, mas possuem o mesmo valor probabilidade. Já as sentenças "a estrela da manhã implodiu" e " a estrela da noite implodiu" não possuem o mesmo conteúdo informacional, embora possam conter o mesmo valor probabilidade em uma dada situação. Este é o caso no qual "estrela da manhã" e "estrela da noite" se referem ao mesmo objeto, fazendo com que a sentença seja associada ao mesmo evento.

Do postulado (P1) temos, obviamente, que $P(\varphi) \neq P(\psi) \Rightarrow I_S(\varphi)$ não é a mesma $I_S(\psi)$. Isso também parece bastante razoável. Se, de acordo com uma dada situação, a chance de ocorrer aquilo que está sendo dito por uma sentença difere da chance de ocorrência daquilo que está sendo dito por outra sentença, então elas não podem estar falando da mesma coisa. O seu conteúdo informacional deve ser distinto. Neste sentido, dado que as sentenças "caiu par" e "caiu primo" não possuem a mesma probabilidade de ocorrência, o seu conteúdo informacional deve ser distinto, como realmente ocorre.

O postulado (P2) explicita que não vale a recíproca do postulado (P1): sentenças com probabilidade igual nem sempre dizem o mesmo. Isso pode ser ilustrado a partir do exemplo do jogo de dado, em que a fórmula A_1 é associada ao evento "par" e a fórmula A_2 é associada ao evento "ímpar": $P(\varphi) = P(\psi)$, mas $I_S(\varphi)$ não é a mesma $I_S(\psi)$. Embora a probabilidade de "caiu par" seja igual à probabilidade de "caiu ímpar" no jogo de um dado não viciado, o conteúdo informacional de cada sentença é bastante distinto.

Os seguintes resultados provados a partir de $S_{\mathbb{P}}$ contrariam os dois postulados da informação semântica expostos acima:

a: $I(\varphi) = I(\psi) \not\Rightarrow P(\varphi) = P(\psi)$ – Recíproca do Teorema 6.2.4c.

b: $P(\varphi) = P(\psi) \Rightarrow I(\varphi) = I(\psi)$ – Teorema 6.2.4c.

Se aceitarmos os dois postulados da informação semântica expostos acima, mostramos que $S_{\mathbb{P}}$ não captura o conteúdo informacional das sentenças. Ou seja, a informação à qual nos referimos em $S_{\mathbb{P}}$ não diz respeito ao conteúdo informacional das sentenças, mas apenas à sua quantidade de informação.

(CF9) *LSC* não captura a noção de informação tal como definida na TMC (Teorema 6.2.5a).

Ao definir a implicação a partir da disjunção e da negação, *LSC* identifica, indevidamente, o valor informacional de certas fórmulas. No Exemplo 6.2.6 também mostramos que a implicação material probabilística tampouco captura a noção de quantidade de informação.

(CF10) A consequência lógica probabilística e a consequência lógica veritativo-funcional são satisfeitas pelo mesmo conjunto de objetos (Teorema 6.3.7d).

A equivalência entre as duas definições de consequência lógica mostra que elas são satisfeitas pelo mesmo conjunto de objetos. Além disso, juntamente com o Teorema 6.3.7c, esta equivalência mostra que, embora a consequência lógica semântica veritativo-funcional seja definida unicamente em termos de valores de verdade, o que está sendo dito na conclusão de uma consequência lógica veritativo-funcional está garantido pelo que está sendo dito nas suas premissas.

O Teorema 6.3.7b mostra que a consequência lógica probabilística pode ser definida a partir da noção de modelo, como feito usualmente na consequência lógica veritativo-funcional: uma fórmula φ é consequência lógica de um conjunto Γ de fórmulas sse φ é válida em todo modelo de Γ.

Apesar destas semelhanças, as perspectivas semânticas probabilística e veritativo-funcional apresentam algumas características próprias, como mostrado no comentário sobre o Teorema 6.3.7.

O Teorema da Compacidade também pode ser mostrado para a consequência lógica probabilística, conforme dito nos comentários sobre o Teorema 6.3.7e. A rigor, o conjunto de premissas de uma consequência lógica probabilística é finito. Isso não impede que, em algumas ocasiões, possamos garantir que uma fórmula é consequência lógica probabilística de um conjunto infinito de fórmulas. Por exemplo: pelo Teorema 6.3.10b, temos que $\Gamma \subseteq \Delta$ e $\Gamma \vDash_{\mathbb{P}} \varphi \Rightarrow \Delta \vDash_{\mathbb{P}} \varphi$. Assim, se $\Gamma \vDash_{\mathbb{P}} \varphi$, então $\Delta \vDash_{\mathbb{P}} \varphi$, ainda que Δ seja infinito, se ele contiver Γ. Claramente, a demonstração da relação de consequência lógica probabilística neste caso depende diretamente do conjunto finito e não do conjunto infinito.

Há situações em que poderíamos determinar se uma fórmula é consequência lógica probabilística de um conjunto infinito Δ de fórmulas analisando diretamente o conjunto Δ. Seria o caso, por exemplo, de um conjunto contraditório, como na Definição 6.3.2d: seja $\Delta = \{\varphi, \neg\varphi, \varphi_1, \varphi_2, ...\}$. Pelo Teorema 5.4.10k, temos que $P((\varphi \wedge \neg\varphi) \wedge \varphi_i) = 0$. O valor da fórmula toda, neste caso, independe do valor de φ_i ou da sua constituição. Assim, poderíamos estender este resultado, considerando que $P((\varphi \wedge \neg\varphi) \wedge \varphi_1 \wedge \varphi_2 \wedge ...) = 0$. Consequentemente, $\Delta \vDash_{\mathbb{P}} \psi$.

O último caso acima mostra ainda a existência de ocasiões em que certas premissas podem ser desnecessárias para a prova de uma conclusão. A eliminação delas do conjunto de premissas manteria a relação de consequência lógica probabilística entre o novo conjunto, estritamente contido no original, e a conclusão. No exemplo acima, cada fórmula $\varphi_i \in \Delta$, para $i \geq 1$, pode ser descartada de Δ, sem com isso alterar a relação de consequência lógica probabilística, dado que $\varphi \wedge \neg\varphi \vDash_{\mathbb{P}} \psi$. Outro exemplo, desta vez de um conjunto satisfatível de fórmulas, seria o seguinte: $\Delta = \{\varphi \rightarrow \psi, \varphi, \varphi \vee \psi\} \vDash_{\mathbb{P}} \psi$. No entanto, $\Gamma = \{\varphi \rightarrow \psi, \varphi\} \vDash_{\mathbb{P}} \psi$, de modo que $\Gamma \subset \Delta$. A fórmula disjuntiva em questão é totalmente desnecessária para a garantia da conclusão.

(CF11) A consequência lógica probabilística é uma consequência lógica Tarskiana (Teorema 6.3.10).

(CF12) As fórmulas informacionalmente vazias são consequência lógica informacional do conjunto vazio (Teorema 6.4.7a).

Isso quer dizer que as fórmulas probabilisticamente válidas e contraditórias são auto-sustentadas. Este resultado ilustra uma primeira diferença entre as versões veritativo-funcional e probabilística com a versão informacional da consequência lógica. Naquelas duas versões, em geral, uma contradição não é consequência lógica de um dado conjunto de premissas.

(CF13) As fórmulas que são consequência lógica informacional de um dado conjunto de fórmulas cuja quantidade de informação é sempre nula são consequência lógica informacional do conjunto vazio (Teorema 6.4.7c).

Este resultado também ilustra uma característica própria da perspectiva informacional da consequência lógica. Dele, resulta que, se uma fórmula é consequência lógica de um conjunto contraditório de fórmulas, ela é consequência lógica informacional do conjunto vazio. Isso, em geral, não vale para as perspectivas veritativo-funcional e probabilística da consequência lógica.

(CF14) Algumas regras de inferência dos sistemas formais lógicos clássicos não possuem validade geral na perspectiva informacional da consequência lógica (Proposição 6.4.8).

Dentre estas regras encontram-se as regras da expansão, do corte e *modus ponens*. Os casos em que o conjunto de premissas possui informação nula configuram exemplo em que a conclusão pode ser mais informativa do que o conjunto de premissas. No entanto, quando a quantidade de informação do conjunto de premissas for maior do que zero, a quantidade de informação da conclusão é sempre menor.

Deste resultado também podemos concluir que a perspectiva informacional da consequência lógica não é equivalente às outras duas perspectivas investigadas neste trabalho. Há fórmulas que são consequência lógica probabilística e, portanto, veritativo-funcional, de um dado conjunto de fórmulas, mas não são consequência lógica informacional. Por outro lado, como já mostrado a partir do Teorema 6.4.7d ou do Corolário 6.4.10c, há fórmulas que são consequência lógica informacional de um dado conjunto de fórmulas, mas não o são do ponto de vista probabilístico ou veritativo-funcional.

(CF15) Uma fórmula é consequência lógica informacional de um dado conjunto de fórmulas sse o dado conjunto de premissas é probabilisticamente equivalente à conclusão ou a conclusão é informacionalmente nula (Teorema 6.4.9).

Deste resultado mostramos que uma fórmula probabilisticamente contraditória pode ser consequência informacional de um dado conjunto de fórmulas informativas. Por exemplo: $\varphi \Vdash \neg\varphi \wedge \varphi$.

Se, por um lado, a consequência lógica informacional se distingue da consequência lógica tradicional, dedutiva clássica, ela tampouco pode ser considerada uma inferência indutiva. Isto por que estas inferências se caracterizam justamente por permitir que a conclusão possua mais informação do que o seu conjunto de premissas.

(CF16) A Lógica subjacente à consequência lógica informacional é, no mínimo, Paraconsistente, *lato-sensu* (Corolário 6.4.10a).

Do Corolário 6.4.10a concluímos ainda que a Lógica Clássica não é a Lógica subjacente à consequência lógica informacional. Isso porque, nos sistemas formais lógicos clássicos, uma contradição gera qualquer fórmula. Consequentemente, as Lógicas Complementares tampouco podem ser subjacentes à consequência lógica informacional, dado que elas satisfazem os princípios da Lógica Clássica. Restam como candidatas a essa Lógica as Lógicas Heterodoxas, como, por exemplo, os sistemas Intuicionistas, Temporais ou as Paraconsistentes. Do Corolário 6.4.10 obtemos que os sistemas intuicionistas não podem ser e do Teorema 6.4.6d mostramos que tampouco as lógicas temporais. Acreditamos que a melhor candidata a esta lógica seria, no mínimo, Paraconsistente.

Um dos motivos para chegarmos a esta conclusão é que, nas lógicas paraconsistentes, não vale o princípio da explosão: de uma contradição nem sempre se segue qualquer coisa. Uma proposta de trabalho futuro seria analisar as características elementares de um sistema paraconsistente e mostrar que a consequência lógica informacional os satisfaz.

Do Corolário 6.4.10a obtemos que, em sistemas formais lógicos que adotam a consequência lógica informacional, a inconsistência da teoria não implica a sua trivialidade, como ocorre com as perspectivas veritativo-funcional e probabilística. Na versão probabilística, este resultado advém da Definição 6.3.2 e do Teorema 6.3.3, a partir dos quais obtemos que: se $\Gamma \subseteq \text{Teo}(T) = \{\varphi \mid \vdash_T \varphi\}$ então Γ é contraditório $\Rightarrow$ T é inconsistente. Do ponto de vista de sistemas formais lógicos como os clássicos, isso significa que qualquer fórmula seria considerada um teorema da teoria T. Consequentemente, pelo Teorema da Completude para estes sistemas, $\vDash_{\mathcal{V}} \psi$ e $\vDash_{\mathbb{P}} \psi$. No entanto, como na consequência lógica informacional, em geral, $\perp_{\mathbb{P}} \Vdash\!\!\!\!/\; \varphi$, não se pode concluir que a inconsistência de uma teoria implique a sua trivialidade.

O Corolário 6.4.10a reforça que a consequência lógica informacional captura a noção intuitiva de consequência lógica informacional, baseada na noção quantitativa de informação. Ao dizer que “hoje é terça e hoje não é terça” não posso concluir disso nem que hoje é terça nem que hoje não é terça ou qualquer outra sentença, como “a bengala está no canto”. Nas premissas eu nada digo a respeito de bengalas ou qualquer outra coisa que não seja hoje é terça.

(CF17) O correspondente semântico do Teorema da Dedução não é válido na perspectiva informacional da consequência lógica (Teorema 6.4.11c).

No Teorema 6.4.11 também mostramos outros resultados expressivos, próprios da perspectiva informacional: uma biimplicação entre duas fórmulas informacionalmente vazia nem sempre implica que as duas fórmulas sejam equivalentes; dizer que duas fórmulas são informacionalmente equivalentes não garante que ambas são informacionalmente vazias em toda situação; dizer que uma fórmula é consequência de um conjunto de fórmulas não significa dizer que o evento associado ao conjunto de premissas está sempre contido no evento associado à conclusão.

(CF18) A consequência lógica informacional não é uma consequência lógica Tarskiana (Teorema 6.4.12)

Grande parte das características próprias da perspectiva informacional da consequência lógica, como a que se refere à não-monotonicidade, é oriunda dos casos em que a quantidade de informação é nula. A diferença específica desta perspectiva encontra-se nos casos em que se trata de valores probabilidades extremos, seja em situações particulares, seja em todas as situações.

Embora seja esta a diferença básica, não podemos considerá-la uma pequena diferença. Ela produz resultados que podem ser considerados discrepantes, quando comparados às perspectivas tradicionais da consequência lógica. Dentre elas, relembramos que algumas regras de inferência não constituem argumento válido ou que nem toda fórmula é consequência lógica informacional de um conjunto contraditório de fórmulas. Isto mostra ainda que a consequência lógica, quando analisada sob a ótica informacional em questão, deixa de ser Tarskiana, e que a lógica subjacente a esta perspectiva não é clássica, mas sim, no mínimo, paraconsistente.

(CF19) O conjunto de premissas de uma inferência informacional é sempre finito e o espaço amostral é constituído por um conjunto finito de eventos.

Estas duas características representam restrições sérias em nossa proposta. A primeira restrição indica que não podemos lidar com argumentos com um conjunto de premissas potencialmente infinito, como os ω-argumentos, discutidos na Seção 2.4.

A segunda restrição limita os modelos possíveis para uma linguagem de um sistema formal. Uma proposta de trabalho futuro seria tratar da consequência lógica informacional a partir de uma definição que envolva um espaço de probabilidades infinito. A consideração de um espaço de probabilidade infinito também tornaria viável uma análise da consequência lógica informacional para uma linguagem de uma teoria de primeira ordem, o que não fizemos neste trabalho.

(CF20) A consequência lógica informacional não satisfaz a característica da necessidade.

Parece que a consequência lógica informacional satisfaz a característica da necessidade, quando diz, para toda situação.

Certamente sentenças probabilisticamente válidas como "chove ou não chove" são necessárias por estarmos pressupondo que, em cada instante de tempo, alguma coisa acontece e que, além disso, o experimento aleatório é estático. Ou seja, aceitamos experimentos em que determinamos tudo o que pode ocorrer, que alguma coisa sempre acontece e que o experimento não se altera com o tempo. Por isso, talvez essa concepção de validade talvez não satisfaça a característica da anterioridade ou da necessidade da consequência lógica.

A Definição 6.3.4 expressa que a consequência lógica probabilística é estabelecida a partir de uma relação entre um conjunto de premissas e uma conclusão. O valor probabilidade daquele é sempre menor que o valor probabilidade deste. A chance de ocorrer tudo o que está sendo dito pelas premissas é sempre menor do que a chance de ocorrer o que está sendo dito na conclusão. Em outras palavras, quando uma fórmula φ é consequência lógica probabilística de um conjunto Γ de fórmulas, é impossível a probabilidade de Γ ser maior do que a probabilidade de φ. Isso jamais pode ocorrer, neste caso, em qualquer situação imaginável. Esse poderia ser um fator indicativo de que a consequência lógica probabilística satisfaz a característica da necessidade da consequência lógica, tal como descrito na Seção 1.2.

Como visto, nenhuma das perspectivas analisadas satisfazem as características da consequência lógica enunciadas no primeiro capítulo. Isso pode indicar duas hipóteses: ou estas características não são adequadas ou as concepções sugeridas não são satisfatórias.

Parece-nos que as três características são bons indicadores de uma relação de consequência lógica razoável, o que nos faz desconsiderar a primeira hipótese. Ao mesmo tempo, as diferentes concepções de consequência lógica podem ser adequadas se considerarmos o contexto ou objetivo para o qual cada uma delas é construída. Isso, de alguma forma, nos faz desconsiderar a segunda hipótese.

Talvez seja impossível apresentar uma definição de uma consequência lógica absolutamente imune a influências valorativas, tais como epistemológicas ou conceituais, ou a dadas concepções de racionalidade. Disso não resulta um ceticismo, relativismo ou utilitarismo em relação a esta noção. A aceitabilidade de uma definição de consequência lógica pode estar em um equilíbrio entre a satisfação de características como as apontadas no primeiro capítulo e os objetivos ou o contexto em que cada definição é sugerida.

Referências Bibliográficas

ANDERSON, A.R; BELNAP, N.D. *Entailment*: the logic of relevance and necessity. Princeton: Princeton University Press, 1975.

ARISTOTLE. *The complete works of Aristotle*. Ed. Jonathan Barnes. Princeton: Princeton University Press, 1984. v. 1.

BARWISE, J; ETCHEMENDY, J. *The liar*: an essay on truth and circularity. New York: Oxford University Press, 1987.

BAYS, T. On Tarski on models. *The Journal of Symbolic Logic*, v. 66, n. 4, p. 1701-26, 2001.

BLAHUT, R. *Principles and practice of information theory*. Amsterdam: Addison-Wesley Publisher, 1988.

BLANCHETTE, P. A. Logical consequence. In: GOBLE, L. (Ed.). *The Blackwell guide to philosophical logic*. Malden, Mass.: Blackwell Publishers, 2001. p. 115-135.

BOLZANO, B. *Theory of science.* Oxford: Basil Blackwell, 1973.

BRESCIANI FILHO, E; D'OTTAVIANO, I. M. L. Conceitos básicos de sistêmica. In: D'OTTAVIANO, I. M. L; GONZALES, M. E. Q. (Orgs.). *Auto-organização*: estudos interdisciplinares. Campinas: Universidade Estadual de Campinas/ CLE, 1990. p. 283-306. (Coleção CLE, v. 30)

CARNAP R. *The logical foundations of probability.* Chicago: Chicago University Press, 1962.

CARNIELLI, W; EPSTEIN, R. *Computabilidade, funções computáveis, lógica e os fundamentos da matemática.* São Paulo: Editora UNESP, 2006.

CIGNOLI, L. O; D'OTTAVIANO, I. M .L; MUNDICI, D. *Álgebras das lógicas de Łukasiewicz*. Campinas: Universidade Estadual de Campinas/CLE, 1992. p. 65-93. (Coleção CLE, v. 12)

_____. *Algebraic foundations of many-valued reasoning*. Dordrecht: Kluwer Academic, 2000.

COHEN, M.; NAGEL, E. *An introduction to logic and scientific method.* London: Routledge and Kegan Paul, 1949.

CORCORAN, J. Aristotle's demonstrative logic. *History and Philosophy of Logic*, v. 30, n.1, p. 1-20, 2009.

DA COSTA, N. C. A. Nota sobre o conceito de contradição. *Anuário da Sociedade Paranaense de Matemática*, n. 1, p 6-8, 1958.

______. *Sistemas formais inconsistentes*. 1963. Tese (Doutorado) – Universidade Federal do Paraná, Curitiba, 1963.

______. On the theory of inconsistent formal systems. *Notre Dame J. of Formal Logic*, 15, p. 497-510, 1974.

______. *Lógica indutiva e probabilidade*. 2.ed. São Paulo: Ed. HUCITEC, 1993.

______. *Logiques classiques et non classiques*: essai sur les fundements de la logique. Paris: Masson, 1997.

DEBRUN, M. A ideia de auto-organização. In: DEBRUN, M; GONZALES, M. E. Q; PESSOA JR, O. (Orgs.). *Auto-organização*: estudos interdisciplinares. Campinas: Universidade Estadual de Campinas/CLE, 1996a. p. 3-23. (Coleção CLE, v.18)

______. A dinâmica da auto-organização primária. In: DEBRUN, M; GONZALES, M. E. Q; PESSOA JR, O. (Orgs.) *Auto-organização*: estudos interdisciplinares. Campinas: Universidade Estadual de Campinas/CLE, 1996b. p. 25-59. (Coleção CLE, v.18).

DEVLIN, K. *Logic and information.* Cambridge: Cambridge University Press, 1991.

D'OTTAVIANO, I. M. L. Paradoxos auto-referenciais e as lógicas não-clássicas heterodoxas. *Ciência e cultura*, v. 42, n. 2, p. 164-73, 1990.

______. On development of paraconsistent logic and da Costa's work. In: *The Journal of Non-Classical Logic*, v. 7, p. 89-152, 1990a.

______. A lógica clássica e o surgimento das lógicas não-clássicas. In: ÉVORA, F. R. R. (Org.). *Século XIX*: o nascimento da ciência contemporânea. Campinas: Universidade Estadual de Campinas/CLE, 1992. p.65-93. (Coleção CLE, v. 11)

DOYLE, Sir A. C. *Um estudo em vermelho*: uma aventura de Sherlock Holmes. São Paulo: Martin Claret, 2001. (Coleção a obra-prima de cada autor)

DRETSKE, F. *Knowledge and the flow of information*. Oxford: Basil Blackwell, 1981.

EDWARDS, E. *Introdução à teoria da informação.* São Paulo: Cultrix, 1971.

EPSTEIN, I. *Teoria da informação.* São Paulo: Ed. Ática, 1986. (Série Princípios)

ETCHEMENDY, J. *The concept of logical consequence.* Stanford: CSLI Publications, 1999.

FEITOSA, H. A. *Traduções conservativas*. 1998. 161 f. Tese (Doutorado em Filosofia) – Instituto de Filosofia e Ciências Humanas, Universidade Estadual de Campinas, Campinas, 1998.

______; D'OTTAVIANO, I. M. L. Um olhar algébrico sobre as traduções intuicionistas. In: SAUTTER, F. T; FEITOSA, H. A. (Orgs.) *Lógica:* teoria, aplicações e reflexões. Campinas: Universidade Estadual de Campinas/CLE, 2004. p. 59-90. (Coleção CLE, v. 39)

FIELD, H. *Realism, mathematics and modality*. Oxford: Basil Blackwell, 1989.

GÖDEL, K. On formally undecidable propositions of Principia mathematica and related systems I. In: ______. *Collected Works*. Oxford: Oxford University Press, 1986. v. 1: Publications 1929-1936, p. 144-95.

GOMES, F. V. *Entre o racional e o justo:* a lógica e as sentenças judiciais. 2008. 127 f. Dissertação (Mestrado em Filosofia) – Instituto de Filosofia e Ciências Humanas, Universidade Estadual de Campinas, Campinas, 2008.

GÓMEZ-TORRENTE, M. Tarski on logical consequence. *Notre Dame Journal of Formal Logic*, v. 25, p. 617-77, 1996.

GONZALEZ, M. E. Um estudo cognitivo-informacional das representações mentais. In: ABRANTES, P. (Org.). *Epistemologia e cognição.* Brasília: Ed. UnB, 1994. p. 127-146.

______. Informação e cognição: uma proposta de (dis)solução do problema mente-corpo. In: ENCONTRO BRASILEIRO/INTERNACIONAL DE CIÊNCIAS COGNITIVAS, 2., 1996, Campos dos Goytacazes. *Anais*... Campos de Goytacazes: Universidade Estadual do Norte Fluminense, 1996. p. 53-60.

HAACK, S. *Filosofia das lógicas*. Tradução de C. A. Mortari e L. H. A. Dutra. São Paulo: Editora UNESP, 2002.

HANSON, W. H. The concept of logical consequence. *The Philosophical Review*, v. 106, n.3, p. 365-409, July, 1997.

HARTLEY, R. Transmission of information. *Bell System Technical Journal*, v. 7, p. 535-563, 1927.

HERSHBERGER, W. *Principles of communication systems.* New York: Prentice-Hall, 1955.

KOLMOGOROV, A. N. O principe tertium non datur. *Matematiceskij Sbornik*, v. 32, p. 646–667, 1925.

KRIPKE, S. Semantical considerations on modal logic. *Acta Philosophica Fennica*, v. 16, p. 83-94, 1963.

______. *Naming and necessity*. Harvard: Harvard University Press, 1980.

LE COADIC, Y. F. *A ciência da informação.* Tradução de Maria Yeda F. S. de F. Gomes. Brasília: Briquet de Lemos, 1996.

LEIBNIZ, G. W. *Philosophical Essays*. Ed. e tradução Roger Ariew e Daniel Garber. Indianápolis: Hackett Publishing, 1989.

LEWIS, C. I. *A survey of symbolic logic.* California: Berkeley University of California Press, 1918.

______; LANGFORD, C.H. *Symbolic logic*. New York: The century company, 1932

ŁUKASIEWICZ, J. On three-valued logic. In: BORKOWSKI, L. (Ed.) *Selected works by Jan Łukasiewicz*. Amsterdam: North–Holland, 1970. p. 87–88.

______. *Selected works of J. Łukasiewicz.* Ed. L. Borkowski. Amsterdam: North-Holland, 1970.

MATES, B. *Lógica Elementar.* Tradução de L. Hegenberg e O. S. Mota. São Paulo: Companhia Editora Nacional/Editora da Universidade de São Paulo, 1968.

MENDELSON, E. *Introduction to logic*. New York: D. Van Nostrand Company, 1964.

MONK, R. *Bertrand Russell*. Tradução de Luiz H. Dutra. São Paulo: Editora UNESP, 2000. (Coleção grandes filósofos)

MORTARI, C. *Introdução à lógica*. São Paulo: Editora UNESP, 2001.

NAGEL, E.; NEWMAN, J. *A prova de Gödel*. Tradução de GUINSBURG, G. 2.ed. São Paulo: Ed. Perspectiva, 2003.

NETTO, J. T. C. *Semiótica, informação e comunicação.* São Paulo: Ed. Perspectiva, 2001. (Coleção debates).

ORLOV, I. The logic of compatibility of propositions. *Matematicheskii Sbornik*, v. 35, n. 3/4, p. 263-286, 1928.

PEIRCE, C.S. *The Collected Papers of Charles Sanders Peirce*. HARTSHORNE, C., WEISS, P. e BURKS, A. (org.) Cambridge, Massachusetts: Harvard University Press, 1958; 8 vols.

PEREIRA, A; GONZALEZ, M. E. Q. Informação, organização e linguagem. In: ÉVORA, F. R. R. (Ed.). *Espaço e tempo.* Campinas: Universidade Estadual de Campinas/CLE, 1995. p. 255-90. (Coleção CLE, v. 15)

PIGNATARI, D. *Informação, linguagem, comunicação.* São Paulo: Ed. Perspectiva, 1968.

QUINE, W. *El sentido de la nueva lógica.* Tradução de Mario Bunge. Buenos Aires: Nueva Visión, 1958.

______. *Los metodos de la logica*. Barcelona: Ariel Ed, 1967.

______. *Philosophy of logic*. England Cliffs: Prentice Hall, 1970.

RAY, G. Logical Consequence: a defense of Tarski. *Journal of Philosophical Logic,* v. 25, n. 6, p. 617-677, 1996.

READ, S. *Thinking about logic*: an introduction to philosophy of logic. Oxford: Oxford University Press, 1994.

ROMAN, S. *Coding and information theory.* New York*:* Springer, 1992.

RUYER, R. *A cibernética e a origem da informação.* Rio de Janeiro: Paz e Terra, 1972.

SHANNON, C.; WEAVER, W. *The mathematical theory of information.* Urbana: University of Illinois Press, 1949.

SHER, G. Y. Did Tarski commit "Tarski Fallacy"? *The Journal of Symbolic Logic*, v. 61, n. 2, p. 653-686, 1996.

SHOENFIELD, J. *Mathematical logic.* Reading, Mass.: Addison Wesley Publishing Company, 1967.

STONIER, T. *Information and the internal structure of the universe.* London: Springer-Verlag, 1990.

TARSKI, A. On some fundamental concepts of metamathematics. In: ______. *Logic, semantics, metamathematics*. Oxford: Clarendon Press, 1956a. p. 30-37.

______. Fundamental concepts of methodology of deductive sciences. In: ______. *Logic, semantics, metamathematics*. Oxford: Clarendon Press, 1956b. p. 60-109.

______. Concept of truth in formalized languages. In: ______. *Logic, semantics, metamathematics*. Oxford: Clarendon Press, 1956c. p. 152-278.

______. On the concept of logical consequence. In: ______. *Logic, semantics, metamathematics*. Oxford: Clarendon Press, 1956d. p. 409-420.

______. The semantic conception of truth*:* and the foundations of semantics. *Philosophy and Phenomenological Research*, v. 4, n. 3, p. 341-76, 1944.

______. *Introduction to logic and the methodology of deductive sciences*. New York: Oxford University Press, 1965.

WIENER, N. *O conceito de informação na ciência contemporânea*. Rio de Janeiro: Paz e Terra, 1970. (Ciência e informação, v.2)

YVES, F. *A ciência da informação.* Brasília: Briquet de Lemos Livros, 1996.