ПРОБЛЕМА ПРИНЯТИЯ РЕШЕНИЙ НА ОСНОВЕ ЛОГИК С ВЕКТОРНОЙ СЕМАНТИКОЙ

Л.В.Аршинский

Abstract. *In this article the problem of logical inference by means of one class of vector semantics logics is discussed. In these logics a truth (faithfulness) is described as a ⟨True; False⟩-vector with two components. These components are independent from each other. The problems of two forms of conjunction, disjunction and negation, the problems of inference rules and interval inference are considered*.

Особенностью некоторых задач принятия решений является плохая формализованность соответствующей предметной области, информация о которой, кроме того, порой имеет неполный и противоречивый характер. При этом противоречие может выражаться во взаимоисключающих сведениях, поступающих от разных источников, принципиальной полисемичности используемых суждений, а неполнота – в отсутствии или недостатке необходимых сведений. Одним из распространенных подходов к принятию решений в подобных предметных областях является использование методов и приемов теории экспертных систем, что требует привлечения к обработке данных средств логического вывода. Однако классический вывод, основанный на точном знании об Истине либо Лжи суждения, здесь оказывается неприемлем, ибо имеет весьма ограниченные возможности по обработке неполных и противоречивых данных. Причина понятна: он основан на законах противоречия и исключения третьего, которые плохо отвечают данной ситуации. В этих случаях традиционно используются различные варианты неклассического логического вывода, основанные на многозначных или нечетких логиках и их аналогах: нечеткий, вероятностный вывод и т.п., поскольку в них провозглашается отказ от указанных логических принципов. Во многих случаях этого оказывается вполне достаточно. Однако, не смотря на то, что подобные логики позволяют рассуждать в указанных предметных областях, они все же не целиком свободны от упомянутых ограничений. В частности, в многозначных логиках суждениям запрещено иметь более одного значения истинности, в них не допускаются значения истинности сверх оговоренного набора таких значений (принцип исключения n-го); в нечетких логиках фиксируется взаимосвязь между Истиной и Ложью суждения и т.п. В докладе обсуждается еще один подход к описанию истинности суждений о неполных и противоречивых предметных областях. Он основан на представлении истинности вектором ⟨Истина; Ложь⟩, где оба компонента (*аспекта истинности*) принимают значения из интервала [0, 1] [1]. Это делает Истину и Ложь не зависящими друг от друга, а также естественно допускает наличие иных аспектов, кроме Истины и Лжи (без изменения основного формализма). Нечеткие, а также некоторые конечнозначные логики являются частными случаями такого взгляда. Его частным случаем является и классическая логика. Обсуждаются вопросы логического вывода для такого представления.

Идея независимости Истины и Лжи не является новой. В 1947 г. в работе Д.Нельсона [2] был предложен вариант конструктивной логики, в которой отсутствует асимметрия между Истиной и Ложью. Ложь в ней требовала таких же доказательных обоснований, как и Истина. В 1966 году в диссертации американского логика Дж.М.Данна (см. [3]) была предложена семантика, допускающая для суждений помимо обычных значений истинности: Истины либо Лжи, также отсутствие того и другого, и их совместную реализацию. Формально это выражалось в том, что в качестве значений истинности рассматривались все подмножества двухэлементного множества {Истина, Ложь}: {{Истина, Ложь}, Истина, Ложь, $\varnothing$}. Содержательно совместную реализацию Истины и Лжи предлагалось интерпретировать как наличие противоречивых оснований для оценки истинности, а $\varnothing$ – как их полное отсутствие. Точка зрения на возможность совместной реализации Истины и Лжи, либо их одновременного отсутствия, была развита в работе Н.Белнапа в 1977 г. в связи с особенностями компьютерной обработки данных (данные подтверждают, отрицают, противоречат друг другу или полностью отсутствуют) [4]. Похожих взглядов на истинность будем придерживаться и мы с тем отличием, что при векторном описании истинности Истина и Ложь рассматриваются не как значения (элементы некоторого «семантического множества»), а как наименования для *категорий* (аспектов) истинности, каждая из которых в свою очередь может быть выражена количественным образом независимо от другой (без ограничения общности можно считать, что их количественные значения принадлежат интервалу $[0,1]$). Дополнительным, хотя возможно и не столь принципиальным отличием рассматриваемой семантики от семантики Данна является то, что она может существовать и в предположении, что возможны еще какие-то аспекты истинности кроме Истины и Лжи. Все «иные» истинностные категории (компоненты вектора истинности) просто дают нулевую проекцию на «плоскость» Истина-Ложь (точнее, на квадрат $[0,1]\times[0,1]$, содержащий все возможные значения истинности в предлагаемой семантике). С формальной точки зрения существование «высших» логических измерений ничего принципиально нового в рассматриваемую логику не вносит. В некотором смысле такой взгляд на истинность есть «синтез» семантики Данна и «нечеткой» семантики, в которой

$$\textit{мера Лжи} = 1 - \textit{мера Истины}. \qquad (1)$$

Логики, основанные на такой семантике, были названы *векторными нестрогими* или V^{TF}-*логиками* (векторные логики с True/False – семантикой) [1]. Впрочем, их с достаточным основанием можно также назвать и *сверхнечеткими логиками* (*super-fuzzy logics*). Термин «сверхнечеткость» оправдан тем, что в отличие от обычных нечетких логик Истина и Ложь здесь не зависят друг от друга.

Принятие решений на основе описанной техники предполагает приписывание исходных значений Истины и Лжи обрабатываемым суждениям. Основанием для этого могут служить экспертные оценки, свидетельские или инструментальные показания, достоверно известные или сомнительные факты с соответствующими мерами доверия к этим фактам или показаниям и т.п. Принципиальной особенность рассматриваемой техники является то, что при ее использовании нет необходимости вычислять «средневзвешенное» взаимоисключающих свидетельств, выдвигать гипотезы об истинности в случае отсутствия или слабой силы свидетельств, огрублять результат до какой-то пороговой величины и т.д., подобно тому, как это делается, например, в нечетких логиках, или когда мы хотим воспользоваться логикой конечнозначной. Оценки «за» (Истина) и «против» (Ложь) здесь накапливаются отдельно. Они определяются, если можно так выразиться, «конструктивным» образом и не «довычисляются» на основе каких-либо «дополнительных» гипотез вроде (1).

Качественно новой особенностью обсуждаемых логик является наличие двух форм конъюнкции, дизъюнкции и отрицания. Так, первая форма конъюнкции и дизъюнкции может быть задана соотношениями [1]:

$$[a \mathbin{\&} b] = \langle a^+ \bullet b^+;\, a^- \oplus b^- \rangle \text{ и } [a \vee b] = \langle a^+ \oplus b^+;\, a^- \bullet b^- \rangle,$$

где $\bullet$ и $\oplus$ – т.н. операции композиционного умножения и сложения: триангулированные норма и ко-норма [5] с дополнительно введенными аксиомами:

$$(1 - x) \bullet (1 - y) = 1 - x \oplus y \text{ и } (1 - x) \oplus (1 - y) = 1 - x \bullet y,$$

справедливость которых обеспечивает переход от рассматриваемых к нечетким логикам в случае введения связи

$$a^- = 1 - a^+.$$

Числа a^+, a^-, b^+, b^- – меры (степени) Истины и Лжи суждений a и b соответственно. Запись $[a]$ следует читать как «истинность a». Собственно аксиомы триангулированных нормы и ко-нормы в данной записи выглядят следующим образом [5]:

1) $x \bullet y = y \bullet x$; $\quad x \oplus y = y \oplus x$;

2) $x' \bullet y \le x'' \bullet y$, при $x' \le x''$; $\quad x' \oplus y \le x'' \oplus y$, при $x' \le x''$;

3) $x \bullet 1 = x$; $\quad x \oplus 0 = x$;

4) $x \bullet (y \bullet z) = (x \bullet y) \bullet z$; $\quad x \oplus (y \oplus z) = (x \oplus y) \oplus z$.

Примерами подобных операций являются известные пары выражений

$$x \bullet y = min(x, y); \qquad x \oplus y = max(x, y);$$

$$x \bullet y = xy; \qquad x \oplus y = x + y - xy;$$

$$x \bullet y = max(0, x + y - 1); \qquad x \oplus y = min(1, x + y).$$

Первые формы конъюнкции и дизъюнкции при переходе к нечетким или классической логикам превращаются в обычные и широко известные операции конъюнкции и дизъюнкции. Вторая форма конъюнкции и дизъюнкции определяется как

$$[a \mathbin{\&_2} b] = \langle a^+ \bullet b^+;\, a^- \bullet b^- \rangle \text{ и } [a \vee_2 b] = \langle a^+ \oplus b^+;\, a^- \oplus b^- \rangle.$$

Она не имеет аналогов ни в нечеткой ни тем более в классической логике. Данные операции характеризуют некую «минимальную» (конъюнкция) и «максимальную» (дизъюнкция) информационную подкрепленность предложений «*a* и *b*» и «*a* или *b*» как по аспекту Истина, так и по аспекту Ложь.

Две формы отрицания соответственно имеют вид:

$$[\neg a] = \langle a^-; a^+\rangle \text{ и } [\sim a] = \langle 1 - a^+; 1 - a^-\rangle.$$

Первая форма отрицания (отрицание в форме перестановки) превращает свидетельства «за» в свидетельства «против» и наоборот и является по своей сути «конструктивным» отрицанием. Вторая форма отрицания – отрицание в форме дополнения – отрицает a ровно в той мере, в какой нет точного знания об a. Оно, если можно так выразиться «гипотетично» по своей сути: мы отрицаем настолько, насколько не знаем. Первое отрицание можно словесно выразить как «точно не–*a*», а второе, как «возможно не–*a*». При переходе к нечетким (и классической) логикам оба этих отрицания сливаются в одно.

Большее количество основных пропозициональных связок, чем в нечетких логиках, создает проблему выбора наиболее адекватных в той или иной ситуации. Частично она была затронута в работе [6]. Ее основным результатом стала рекомендация использовать первую форму конъюнкции, дизъюнкции и отрицания для моделирования естественно-языковых связок «И», «ИЛИ» и «НЕ». Вторую форму конъюнкции и дизъюнкции – для определения того момента, когда соответствующее суждение станет «информационно наполненным» и его можно будет включать в процедуру логического вывода. Для второй формы отрицания в работе [6] места не нашлось, однако есть смысл использовать его как «гипотетическое отрицание», «возможное не-», как предложено выше.

Еще одной проблемой вывода на основе векторных логик является интервальность получающихся значений истинности.

Интервальность следует из правила вывода *modus ponens*, которое в данных логиках имеет две формы: *содержательный modus ponens* (С-МР) и *нестрогий modus ponens* (Н-МР). Причем первое из них имеет две подформы: С-МР и С-МР(2) [1].

<u>Правило С-МР</u>. Если имеется суждение a с вектором истинности $[a] = \langle a^+; a^-\rangle$ и импликация $i = a \to b = \langle i^+; i^-\rangle$, то истинность $[b] = [a \,\&\, i] \div \mathbf{И}$:

$$a, a\to b | \!-\, b\colon [b] = [a \,\&\, i] \div \mathbf{И} = \langle a^+ \bullet i^+; a^- \oplus i^-\rangle \div \langle 1; 0\rangle,$$

где через двоеточие указана схема расчета истинности заключения. Здесь **И** – *строгая Истина* = $\langle 1; 0\rangle$, символ $\div$ означает, что $[b] \in [a^+ \bullet i^+, 1]\times[0, a^- \oplus i^-]$ (левый верхний угол этого прямоугольника задается вектором $\langle a^+\bullet i^+; a^- \oplus i^-\rangle$, а правый нижний вектором $\mathbf{И} = \langle 1; 0\rangle$).

<u>Правило С-МР(2)</u>. Если имеется суждение a с вектором истинности $[a] = \langle a^+; a^-\rangle$ и импликация $i = a \to b = \langle i^+; i^-\rangle$, то истинность $[b] = [a \,\&\, i] \div [\neg a \vee i]$:

$$a, a\to b | \!-\, b\colon [b] = [a \,\&\, i] \div [\neg a \vee i] = \langle a^+ \bullet i^+; a^- \oplus i^-\rangle \div \langle a^- \oplus i^+; a^+ \bullet i^-\rangle.$$

Правило Н-МР в развернутом виде формулируется так.

Правило Н-МР. Если между суждениями a и b существует отношение *правдоподобия* $a < b$ или отношение *доминирования* $a \ll b$, и истинность суждения a равна $[a] = \langle a^+; a^- \rangle$, то в первом случае можно построить вывод:

$$a, a < b |\!- b: [b] = [a] \div \mathbf{И};$$

а во втором вывод:

$$a, a \ll b |\!- b: [b] = [a] \div \mathbf{П},$$

где $\mathbf{П}$ – *полное противоречие* $= \langle 1; 1 \rangle$.

Отношение правдоподобия $a < b$ означает, что $a^+ \le b^+$ и $a^- \ge b^-$, отношение доминирования $a \ll b$, что $a^+ \le b^+$ и $a^- \le b^-$. Правдоподобие означает, что в пользу Истины суждения b получены более сильные свидетельства, чем в пользу Истины a и наоборот в пользу Лжи b – свидетельства менее сильные, чем в пользу Лжи суждения a. Доминирование, – что более сильные свидетельства поступили и в пользу Истины и в пользу Лжи b (суждение b информационно более подкреплено). Похожие отношения возникают и в семантике Данна [7]. Количественное значение истинности b непосредственно следует из этих отношений.

Обратим внимание на то, что для правила С-МР нам нужно знать истинность импликации (причинно-следственной связи), которая должна задаваться экспертно или из каких-либо иных соображений. В случае правила Н-МР этого не требуется. Чтобы вывести заключение, для нас достаточен сам факт такого отношения. Например, $a < a \vee b$, или $a \ll a \vee_2 b$. Зная истинность $[a]$ по Н-МР легко найти истинность $a \vee b$ и $a \vee_2 b$ (точнее, построить «интервалы» значений истинности для них).

Для работы с интервальными значениями истинности можно использовать два подхода.

1. На каждом шаге вывода, когда получаются интервалы значений Истины и Лжи, заменять их «характерными» точечными значениями, например, средним арифметическим или средним геометрическим. Пример такой схемы вывода дан в [1].

2. Работать с интервальными значениями, переходя к характерным только на конечном этапе упорядочения заключений. Упорядочивать их можно сначала, например, по мере достоверности

$$\mu_{д}(b) = b^+ - b^-,$$

а затем – при равных «достоверностях» – по мере определенности:

$$\mu_{о}(b) = b^+ \oplus b^-,$$

или мере избыточности:

$$\mu_{и}(b) = b^+ + b^- - 1.$$

Для работы с интервалами можно использовать интервальное представление операций композиционного умножения и сложения:

Определение 1. Композиционной суммой двух интервалов $[x_1, x_2] \subseteq [0,1]$ и $[y_1, y_2] \subseteq [0,1]$ назовем интервал

$$[x_1, x_2] \oplus [y_1, y_2] = [x_1 \oplus y_1, x_2 \oplus y_2].$$

Определение 2. Композиционным произведением двух интервалов $[x_1, x_2] \subseteq [0,1]$ и $[y_1, y_2] \subseteq [0,1]$ назовем интервал

$$[x_1, x_2] \bullet [y_1, y_2] = [x_1 \bullet y_1, x_2 \bullet y_2].$$

Очевидно, что и $[x_1 \oplus y_1, x_2 \oplus y_2]$ и $[x_1 \bullet y_1, x_2 \bullet y_2]$ также входят в отрезок [0,1].

Вводя частичный порядок между интервалами по принципу

$$[x_1, x_2] \le [y_1, y_2], \text{ если } x_1 \le y_1 \text{ и } x_2 \le y_2$$

легко убедится, что

$$[x_1, x_2] \bullet [y_1, y_2] \le [x_1, x_2], [y_1, y_2] \le [x_1, x_2] \oplus [y_1, y_2].$$

В силу того, что число x можно рассматривать как интервал $[x, x]$, легко видеть также, что в силу перечисленных аксиом

$$x \oplus [y_1, y_2] = [x \oplus y_1, x \oplus y_2];$$

$$x \oplus [0, y] = [x, x \oplus y];$$

$$x \oplus [y, 1] = [x \oplus y, 1].$$

Аналогично,

$$x \bullet [y_1, y_2] = [x \bullet y_1, x \bullet y_2];$$

$$x \bullet [0, y] = [0, x \bullet y];$$

$$x \bullet [y, 1] = [x \bullet y, x].$$

Данные определения композиционной суммы и композиционного произведения для интервалов в целом соответствуют представлениям об операциях интервальной математики [8]. Основываясь на этих операциях, мы и будем строить процедуру накопления свидетельств.

Особенностью присоединенного вывода является возможность накапливать свидетельства в пользу того или иного заключения в случае получения его по нескольким цепочкам вывода. Одним из подходящих приемов такого накопления является накопление по схеме 11- композиции [1], когда, например, в случае получения двух значений вектора истинности:

$$[b]_1 = \langle b_1^+; b_1^- \rangle \text{ и } [b]_2 = \langle b_2^+; b_2^- \rangle.$$

итоговое («накопленное») значение истинности получается по формуле:

$$[b] = \langle b_1^+ \oplus b_2^+; b_1^- \oplus b_2^- \rangle.$$

В интервальной форме оно выглядит так: если

$$[b]_1 = \langle[b_{11}^{+}, b_{12}^{+}]; [b_{11}^{-}, b_{12}^{-}]\rangle \text{ и } [b]_2 = \langle[b_{21}^{+}, b_{22}^{+}]; [b_{21}^{-}, b_{22}^{-}]\rangle.$$

то

$$[b] = \langle[b_{11}^{+} \oplus b_{21}^{+}, b_{12}^{+} \oplus b_{22}^{+}]; [b_{11}^{-} \oplus b_{21}^{-}, b_{12}^{-} \oplus b_{22}^{-}]\rangle$$

(употребление квадратных скобок для обозначения истинности и для обозначения интервалов ясно из контекста).

Пример работы с интервальными значениями истинности для решения задачи о выборе наиболее предпочтительной гипотезы из множества взаимоисключающих гипотез дан в [9].

Литература

1. Многозначные логики с векторной семантикой/ Аршинский Л.В.; ВСИ МВД России.– Иркутск, 2003.– 46 с.: Рус.– Деп. в ВИНИТИ 13.02.03, № 281-В2003.
2. Nelson D. Constructible falsity// Journal of Symbolic Logic. 1949. Vol. 14. P. 16-26.
3. Dunn J.M. The Algebra of Intensional logics. Doctoral Dissertation. University of Pittsburg, Ann Arbor, 1966 (University Microfilms).
4. Belnap. N. A useful four-valued logic// J.M.Dunn and G.Epstein (eds.). Modern Uses of Multiple-Valued Logic. Dordrecht: D.Reidel Publish Co., 1977. P.8-37 (рус. пер. Н.Белнап, Т.Стил. Логика вопросов и ответов. М.: Прогресс, 1981).
5. Menger K. Statistical metrics // Proc. Nat. Acad. Sci. USA 8.– 1942, p.535-537.
6. Аршинский Л.В. Моделирование языковых связок «И», «ИЛИ», «НЕ» средствами векторной логики// Управление в системах: Вестник ИрГТУ. Сер. Кибернетика. – Иркутск: ИрГТУ, 2000. – Вып.3. С.12-18.
7. Шрамко Я.В. Обобщенные истинностные значения: решетки и мультирешетки // Логические исследования. Вып. 9 . – М.: Наука, 2002. – С. 264-291.
8. Moore R.E. Interval Analysis, 1966.
9. Аршинский Л.В. Оценка истинности взаимоисключающих свидетельств средствами векторной логики // Информационные и математические технологии/ Труды Байкальской Всероссийской конференции «Информационные и математические технологии». – Иркутск: ИСЭМ СО РАН, 2004. – С. 188-194.