

УДК 004.82

И.И. Коваленко, А.В. Швед

**НЕКОТОРЫЕ НЕ–ФАКТОРЫ И МЕТОДЫ ИХ МОДЕЛИРОВАНИЯ**

**Введение.** Информационные технологии в последние два десятилетия интенсивно развиваются в рамках сравнительно молодой науки, получившей название «инженерия знаний».

Основу данного научного направления составляют результаты разработок и исследований, связанные с искусственным интеллектом (ИИ): представление знаний и вывод на знаниях; системы искусственного интеллекта (экспертные системы, системы распознавания образов, системы поддержки принятия решений и др.).

Вместе с тем, различные предметные области, знания о которых могут быть положены в основу создания систем искусственного интеллекта, как правило, отрицают традиционные свойства формальных систем: точность, полноту, определенность, корректность, и др. Такие отрицаемые свойства были названы «НЕ–факторы», поскольку каждый из них лексически и содержательно фиксирует учет наших Не–знаний при абстрагировании, переходе к формальным системам и интерпретации выводов, полученных на формальном уровне [1]. Поэтому любая формальная система всегда будет характеризоваться наличием определенных НЕ–факторов.

Данное обстоятельство полностью согласуется с основополагающим принципом неопределенности Вернера фон Гейзенберга и является отражением фундаментальной неопределенности процессов макромира.

В проблематике ИИ учет, анализ и управление НЕ–факторами имеет первостепенное значение, что обусловлено творческим характером задач создания интеллектуальных технологий, которые всегда решаются в условиях противоречивости, неполноты, неточности, неопределенности исходных данных, отношений между ними, операций их обработки (алгоритмов, процессов решения). Вместе с тем, как отмечают авторы работы [1], достаточно часто современные методы нечеткой математики, вероятностно-статистического вывода, байесовские и нейронные сети, генетические алгоритмы и др. используются без должного анализа природы присутствующих НЕ–факторов, что может привести к неадекватным моделям и выводам.

**Постановка задачи.** Целью статьи является рассмотрение подхода, в основе которого лежит систематизация наиболее изученных НЕ–факторов и сравнительный анализ методов их моделирования.

**Изложение основного материала.** Первые НЕ–факторы, получившие название «нечеткость» или «неточность», были определены и изучались в рамках проблематики нечеткой математики, основателем которой является Лотфи Заде [1,2].

Однако, целенаправленные системные исследования НЕ–факторов начались с работ А.С. Нариньяни, в которых введено понятие и дана содержательная их трактовка. Анализ литературных источников [1,3,4,5,6,7] позволяет привести ещё ряд существующих подходов к определению НЕ–факторов.

Так, НЕ–факторы по А.С. Нариньяни [5, 8]: НЕТОЧНОСТЬ, НЕДООПРЕДЕЛЕННОСТЬ, НЕОДНОЗНАЧНОСТЬ, НЕЧЕТКОСТЬ.

НЕ–факторы по В.А. Вагину [4]: ПРОТИВОРЕЧИВОСТЬ, НЕМОНОТОННОСТЬ, НЕТОЧНОСТЬ, НЕОПРЕДЕЛЕННОСТЬ, НЕЧЕТКОСТЬ.

НЕ–факторы по Г.В. Рыбиной [6, 7]: НЕЧЕТКОСТЬ, НЕТОЧНОСТЬ, НЕОПРЕДЕЛЕННОСТЬ, НЕДООПРЕДЕЛЕННОСТЬ.

НЕ–факторы по А.Н. Борисову [3]: НЕИЗВЕСТНОСТЬ, НЕДОСТОВЕРНОСТЬ, НЕОДНОЗНАЧНОСТЬ.

Их рассмотрение приводит к выводу о том, что предложенная совокупность НЕ–факторов достаточно «смутно» отображает принципы их унификации и классификации. Это говорит о том, что тема формализации исследования НЕ–факторов на современном этапе, по-видимому, далека от своего завершения.

Вместе с тем в работах [9, 11,12] делается попытка представить НЕ–факторы, обоснованное существование которых определяется методами их моделирования, базирующимися на традиционных, а также новых развивающихся математических теориях (рис. 1). Рассмотрим их более подробно.

**Неполнота.** Некоторые данные не известны, но вся доступная информация полна и корректна. Например, «Возраст этого индивида неизвестен». Единственным способом сокращения этого незнания является сбор недостающих данных (пополнение информации).

**Неопределённость.** Доступная информация может быть истинной или ложной и оценивается с помощью вероятностных оценок. Например, лицо, которому Вы не доверяете, утверждает, что «Возраст индивида равен 30 годам». Степень доверия этому сообщению может быть выражена с помощью соответствующей вероятностной оценки.

**Неточность (нечеткость).** Имеется информация, достоверность которой не вызывает сомнений, однако она не точна. Например, достоверная информация «Этот индивид молодой» не точна. Анализ такой информации выполняется методами нечеткой логики, нечётких множеств, нечетких отношений [2].

Однако, могут быть ситуации в которых одновременно присутствуют различные формы незнания. Например, лицо, которому Вы не доверяете, говорит «Индивид молодой». В данном случае имеется **комбинация неопределённости и неточности**. Такая комбинация может быть промоделирована с применением методов теории свидетельств Демпстера–Шейфера и теории правдоподобных и парадоксальных рассуждений Дезера–Смарандаке [13, 14, 15].



Рис. 1 Классификация форм незнания и методов их анализа

Рассмотрим приведенные НЕ–факторы с позиции перечисленных теорий, что будет способствовать более четкому их различию.

**Теория вероятностей** оперирует с шансами случайных событий, при этом предполагается, что все события являются хорошо определёнными понятиями. **Неопределённость** здесь связана только с тем, с какими шансами может произойти каждое случайное событие из полной группы таких событий. Для этого необходимо выполнение двух условий: рассматриваются все возможные в данной ситуации события; реализоваться может только одно из событий.

Часто эти два условия формулируются следующим образом: полную группу событий образуют взаимно исключающие и исчерпывающие события. Традиционно вероятностные оценки  $p(e_i)$  случайных событий  $e_i, i = \overline{1, n}$  подчиняются следующей системе аксиом:

1.  $p(e_i) \geq 0, i = \overline{1, n}, 0 \leq p \leq 1$
2.  $P(E) = \sum_{i=1}^n p(e_i), E = \{e_i | i = \overline{1, n}\};$  (1)
3. Если  $E_1 \subseteq E, E_2 \subseteq E$  и  $E_1 \cap E_2 = \emptyset$ , то  $P(E) = p(E_1 \cup E_2) = p(E_1) + p(E_2)$ .

Аксиомы 2 и 3 свидетельствуют об аддитивном характере вероятностных оценок, т.е. в любом случае сумма вероятностей всех случайных событий, образующих полную группу равна  $\sum_{i=1}^n p(e_i) = 1$ .

Следует указать на то, что существует два основных подхода к оцениванию вероятностей событий: объективные оценки вероятностей на основе метода частот и субъективные оценки вероятностей, источниками которых являются эксперты. В первом случае для моделирования неопределенностей могут быть использованы аналитические методы вероятностного вывода, вероятностный вывод на деревьях вероятностей, методы математической статистики, байесовские стратегии и др. Оценки второго вида могут быть использованы для вероятностного вывода на деревьях решений, на сетях уверенностей (сети Байеса),

в задачах классификации (байесовские классификаторы), в задачах принятия решений в условиях риска и др.

**Теория нечетких множеств** предназначена для оперирования нечеткими концепциями, которые лежат в основе формирования множеств элементов. Предполагается, что элементы являются хорошо определенными понятиями. *Неопределенность* (неточность, нечеткость) здесь возникает при попытке отнести элементы к некоторым классам (множествам), поскольку эти классы (множества) являются нечеткими, следовательно, плохо определенными.

Основным математическим аппаратом нечеткой (fuzzy) алгебры и нечеткой логики является лингвистическая переменная (ЛП), значения которой определяются набором вербальных (словесных) характеристик некоторого свойства объекта. Значения ЛП характеризуются так называемыми «нечеткими множествами» (НМ), которые в свою очередь определяются через некоторую базовую числовую шкалу  $A$  и функцию принадлежности  $\mu_A(x)$ ,  $x \in A$ ,  $\mu_A(x) \in [0,1]$ . Таким образом, НМ – это совокупность пар вида  $(x, \mu_A(x))$ , которая определяет субъективную степень уверенности в том, что данное конкретное значение  $A$  соответствует определенному элементу НМ.

Данная теория в качестве «инструментов» моделирования неточностей (нечеткостей) предоставляет методы построения функций принадлежности, формирования баз знаний, анализа нечетких отношений и др. Кроме этого, методы нечетких множеств и нечеткой логики могут быть использованы в задачах классификации, принятия решений, прогнозирования, а также при создании систем искусственного интеллекта, в которых используются сочетания возможностей нечеткой математики с нейронными сетями, генетическими алгоритмами и др.

**Теория свидетельств** Демпстера–Шейфера. Как уже отмечалось выше, теория вероятностей имеет дело с каждым отдельным событием (синглтоном) из полной группы событий и может корректно обращаться с неопределенностями, которые подчиняются аксиомам (1). Вместе с тем, в реальных условиях могут существовать и специфические формы НЕ-факторов, например, комбинация неопределенности и неточности (нечеткости), возникающие в процессе взаимодействия между суждениями экспертов. Формы таких взаимодействий могут иметь различный характер – они могут быть согласованными, совместимыми; могут произвольным образом объединяться и пересекаться. Для моделирования указанных форм взаимодействий может быть использована теория свидетельств Демпстера–Шейфера (ТДШ) [12,13,15], основу которой составляют следующие положения.

Имеется множество элементов  $\Omega = \{\omega_i \mid i = \overline{1, n}\}$ , называемое *основой анализа*. Предполагается, что основа анализа  $\Omega$  представляет собой множество исчерпывающих (всех возможных в данной ситуации) элементов и взаимно исключаемых (уникально определенных и отличных от других) элементов. При этом априори известно, что только единственный элемент  $\omega_0 \in \Omega$  является истинным в каждой конкретной ситуации. Этот элемент называется *действительным миром*.

На основе анализа  $\Omega$  могут быть сформированы произвольные подмножества элементов  $A_i \subseteq \Omega$  при предположениях (свидетельствах), что действительный мир  $\omega_0$  может принадлежать каждому из этих подмножеств.

*Свидетельствами* называют любые источники информации, на основании которых могут быть получены интересующие нас оценки *степеней уверенности*. Подмножества  $A_i$  могут принимать различные формы и взаимодействовать произвольным образом  $A_i = \{\omega_i\}$  – единичное подмножество;  $A_j = \{\omega_2, \omega_3, \omega_4\}$  – многоэлементное подмножество;  $A_i \cup A_j$  – объединение подмножеств  $A_i$  и  $A_j$ ;  $A_i \cap A_j$  – пересечение подмножеств  $A_i$  и  $A_j$  и др.

Подмножества  $A_i$  основы анализа  $\Omega$ , которым эксперт назначил степени уверенности (основные массы вероятности  $m(A_i)$ ) того, что действительный мир может находиться в любом из этих подмножеств, называют *фокальными элементами*.

Следует отметить, что  $m(A_i)$  определяет степень уверенности, отдаваемую  $A_i$ , но никаким подмножествам. Это принципиальное положение теории свидетельств.

Наряду с назначениями масс вероятностей  $m(A_i)$  на подмножества основы анализа  $\Omega$ , концептуальную основу ТДШ составляют такие понятия как *функция уверенности* и *функция правдоподобия*.

Значения первой функции для отдельных подмножеств  $\Omega$  выражают всю степень поддержки, отдаваемую каждому из таких подмножеств. Значения второй функции для каждого из подмножеств  $\Omega$  выражает полную степень потенциальной поддержки, которая может быть отдана каждому из этих подмножеств.

Если на одной и той же основе анализа  $\Omega$  назначения степеней уверенности (основных масс вероятности) выполняют несколько независимых экспертов, то возникает задача комбинирования этих назначений. Это производится на основе различных правил комбинирования свидетельств: Демпстера, Ягера, Инагаки, Дюбуа и Прада и др., которые могут быть использованы в ситуациях наличия специфических

(комбинированных) НЕ–факторов.

**Теория правдоподобных и парадоксальных рассуждений** Дезера–Смарандаке (ТДС) [12, 14] может рассматриваться, как более углубленный вариант ТДШ в том плане, что она может оперировать с более сложными формами НЕ–знания. Например, если элементы основы анализа  $\omega_i$  отражают нечеткие и неопределенные понятия: старость–молодость, цветовую гамму и т.д., то некоторые из элементов могут в значительной степени перекрываться друг другом, поэтому выделить полностью различающиеся  $\omega_i$  не представляется возможным. В основе концепции ТДС лежат понятия *свободной* и *гибридной модели*.

*Свободная модель*, обозначаемая  $M(\Omega)$  рассматривает  $\Omega$  только как основу исчерпывающих элементов  $\omega_i | i = \overline{1, n}$ , которые потенциально могут перекрываться. *Гибридная модель* определяется из свободной модели путем введения некоторых ограничений общности на некоторые подмножества элементов  $A_i$  из множества всех возможных подмножеств  $D^\Omega$ , когда полагается, что  $A_i \neq \emptyset$ . Это объясняется тем, что реальных задачах нет необходимости назначать основные массы вероятностей всем возможным подмножествам  $D^\Omega$ , так как всегда возможны элементы, которые будут взаимно исключающими. В качестве ограничений общности широко применяются: ограничение исключаемости  $\omega_i \cap \dots \cap \omega_k = \emptyset$  и ограничение несуществования  $\omega_i \cup \dots \cup \omega_k = \emptyset$ . Так же, как и в теории свидетельств, в ТДС присутствует правило комбинирования, отражающее конъюнктивный консенсус между различными группами свидетельств. В заключение приведем ряд соображений, характеризующих различия между теорией вероятности, ТДШ и ТДС [12, 14].

Пусть  $\Omega = \{\omega_1, \omega_2\}$  – простейшая основа анализа, состоящая только из двух элементов. Тогда теория вероятностей имеет дело при предположении исключаемости гипотез с основными назначениями вероятностей, такими что  $m(\omega_1) + m(\omega_2) = 1$ ; ТДШ имеет дело при предположении исключаемости и исчерпываемости гипотез с основными назначениями уверенностей, такими что  $m(\omega_1) + m(\omega_2) + m(\omega_1 \cup \omega_2) = 1$ . ТДС имеет дело с единственным предположением исчерпываемости гипотезобобщенными назначениями вероятностей, таким что  $m(\omega_1) + m(\omega_2) + m(\omega_1 \cup \omega_2) + m(\omega_1 \cap \omega_2) = 1$ .

**Выводы.** Рассмотренные теории моделирования НЕ–факторов, несмотря на некоторые кажущиеся аналогии между ними, являются совершенно различными теориями. В основе каждой из них лежит специфический математический аппарат, и они предназначены для моделирования различных НЕ–факторов.

Сходство между этими теориями заключается лишь в способах получения исходной информации. Оценки вероятностей могут быть получены как объективным, так и субъективным (экспертным) путем. В теории нечетких множеств, для оценки степеней принадлежности элементов данному нечеткому множеству, также используются субъективные оценки. В ТДШ и ТДС для получения оценок степеней уверенности используют только субъективные подходы.

Всё это выдвигает условие детального анализа природы НЕ–факторов, что обеспечит правильный выбор методов моделирования, предоставляемых данными теориями.

#### ЛИТЕРАТУРА

1. Валькман Ю.Р., Быков В.С., Рыхальский А.Ю. Моделирование НЕ–факторов – основа интеллектуализации компьютерных технологий. // Системні дослідження та інформаційні технології, 2007. – №1. – С.39 – 61.
2. Заде Л.А. Основы нового подхода к анализу сложных систем и процессов принятия решений // Математика сегодня. – М.: Знание, 1974. – С. 5 – 49.
3. Борисов А.Н., Крумберг О.А., Федоров И.П. Принятие решений на основе нечетких моделей.– Рига: «Зинатне», 1990. – 184 с.
4. Вагин В.Н. Знание в интеллектуальных системах // Новости искусственного интеллекта. – 2002. – №6. – С. 8 – 18.
5. Нариньяни А.С. Неточность как НЕ–фактор. Попытка доформального анализа.– Препринт РосНИИ ИИ, 1994. – № 2. – 34 с.
6. Душкин Р.В., Рыбина Г.В. Об одном подходе к автоматизированному извлечению, представлению и обработке знаний с НЕ–факторами // Изв. РАН. ТиСУ. – 1999. – № 5. – С. 84 – 96.
7. Рыбина Г.В. Модели, методы и программные средства для построения интегрированных экспертных систем: автореферат диссертации на соискание ученой степени д-ра техн. наук: 05.13.11 / Г. В. Рыбина. – М., 2004. – 44 с.
8. Нариньяни А.С. Недоопределенные модели и операции с недоопределенными значениями. – Новосибирск, 1982. – 33 с. (Препр./ АН СССР. Сиб. отд-ние ВЦ; № 400).
9. Коваленко И.И., Швед А.В. Современные методы анализа экспертных оценок // Наукові праці ЧДУ ім. П. Могили, серія „Комп’ютерні технології”, 2012. – вип. 161. – т. 173. – С. 10 – 20.

10. Шокин Ю.И. Интервальный анализ. – Новосибирск: Наука, 1981. – 256 с.
11. Burrus N., Lesage D. Theory of evidence (DRAFT) (Technical Report). Laboratoire de Recherche et Developpement de l'Epita, 2003.
12. Uzga-Rebrovs O. Nenoteiktibu parvaldisana. – Rezekne: RA Izdevnieciba, 2010. – Vol. 3. – 560 lpp.
13. Shafer G. A mathematical theory of evidence. Princeton University Press, 1976. – 297 p.
14. Smarandache F., Dezert J. Representation of DSmT. In Advances and Applications of DSmT for Information Fusion. – American Research Press: Rehoboth, 2004. – Vol. 1 – Pp. 3 – 35.
15. Beynon M.J, Curry B., Morgan P. The Dempster–Shafer theory of evidence: an alternative approach to multicriteria decision modelling // Omega, 2000. – Vol. 28. – № 1. – Pp. 37 – 50.

КОВАЛЕНКО Игорь Иванович – д.т.н., профессор кафедры программного обеспечения автоматизированных систем Национального университета кораблестроения им. Макарова

Научные интересы: методы анализа данных, прикладной системный анализ, теория оптимальных решений, системы поддержки принятия решений

ШВЕД Алена Владимировна – аспирантка кафедры интеллектуальных информационных систем Черноморского государственного университета имени Петра Могилы

Научные интересы: методы анализа данных, математическое моделирование, информационные технологии, системы поддержки принятия решений