

УДК 519.68

DOI: 10.15827/2311-6749.16.4.2

АСПЕКТЫ ОБРАБОТКИ НЕЧЕТКОЙ ТЕМПОРАЛЬНОЙ ИНФОРМАЦИИ*В.В. Фомин, д.т.н., профессор, v_v_fomin@mail.ru;**С.В. Мальцев, аспирант, maltsevsergey@live.com**(РГПУ им. А.И. Герцена, наб. реки Мойки, 48, г. Санкт-Петербург, 191186, Россия)*

В данной работе инициирован процесс рассмотрения и решения проблем хранения и обработки нечеткой темпоральной информации в реляционных БД путем построения новой модели нечетко-темпорального домена и соответствующих типов данных для хранения возможно-временной информации, а также построения методов обработки нечетко-темпоральных отношений между временными точками и/или интервалами с реализацией соответствующих процедур и функций.

Ключевые слова: нечеткая темпоральная информация, темпоральные БД, нечетко-темпоральный домен, нечетко-темпоральные отношения.

В отличие от классической реляционной модели данных применение темпорального подхода при проектировании реляционной БД позволяет использовать дополнительные возможности для хранения не только информации об объектах или фактах совершенных событий, но и о состояниях объектов, которые они принимают в различных временных измерениях. Для любого объекта, созданного в один момент времени и закончившего свое существование в другой, темпоральная поддержка позволяет сохранять все его состояния на временном интервале по временной шкале.

Тем не менее временные показатели и понятия могут содержать неточность, неопределенность, которая в одних случаях может быть приемлемой для удовлетворения требований, а в других необходимой для принятия решений.

Необходимость использования нечеткой логики в обработке темпоральной информации обусловлена возможным наличием неточности в исходных данных, где периоды начала/завершения действия интервалов или событий могут быть неточно определены. Рассмотрим два интервала: интервал I между моментами времени S_i и E_i (рис. 1, #1) и J между моментами времени S_j и E_j (рис. 1, #2). Классическая функция пересечения [1] интервалов I и J определяется как $I \text{ overlaps } J$ при условии, что $S_j < S_i \wedge E_j < E_i \wedge E_j > S_i$. Результатом выполнения классической функции пересечения для интервалов I и J в данном примере будет значение FALSE (рис. 1, пунктир на временных отрезках #1 и #2).

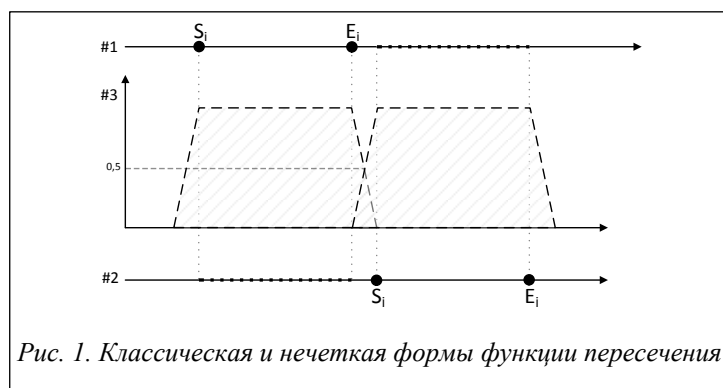


Рис. 1. Классическая и нечеткая формы функции пересечения

Используя нечеткую форму функции пересечения двух интервалов, можно рассчитать, с какой степенью возможности интервал I пересекает интервал J (рис. 1, #3). Если классическая версия функции пересечения интервалов I и J возвратит пользователю ответ «Нет, интервалы не пересекаются», то нечеткая версия функции пересечения возвратит «Пересекаются, со степенью возможности 0.5».

Моделирование темпоральности и нечеткости в БД

В темпоральной модели данных можно выделить следующие типы нечеткости [2].

Неопределенность. Какая-либо часть темпоральной информации неизвестна или определена неточно. К примеру, дата рождения Сократа «около 469 год д.н.э.».

Субъективность, двусмысленность. Темпоральные события или периоды могут быть субъективно или неоднозначно определены. К примеру, «эпоха нового времени» или в сфере грузоперевозки «даты начала или завершения транспортировки груза».

Неясность, расплывчатость. События могут быть определены с различной гранулярностью или нечеткостью. К примеру, «доставка груза в течение дня».

Темпоральная модель данных с использованием вышеописанных типов нечеткости отражена на рисунке 2.

Фактом можно называть утверждение о событии или состоянии действительности, которое может быть проверено, верифицировано. Факты в явном виде могут не иметь никакой принадлежности

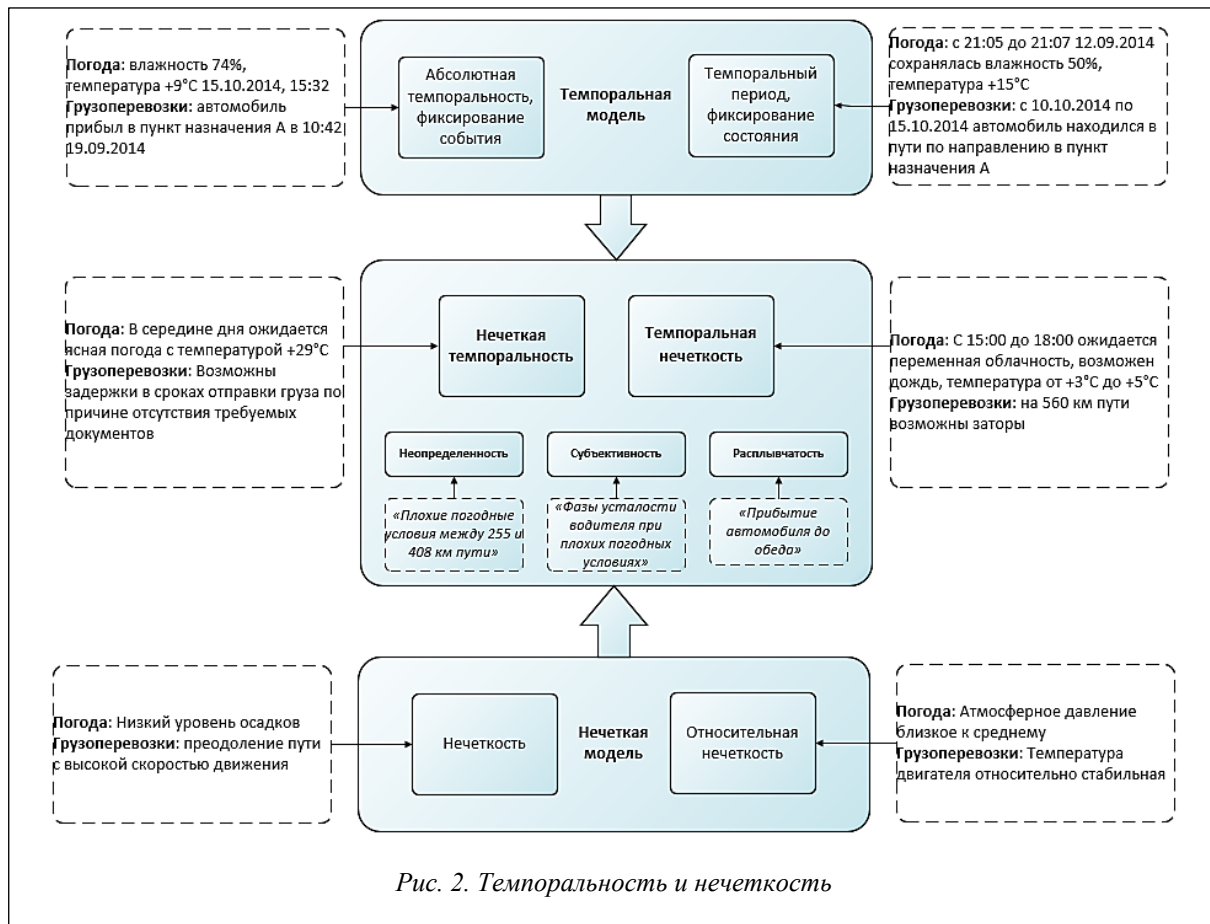


Рис. 2. Темпоральность и нечеткость

к временным показателям, но в неявном виде всегда существует и может быть определена их зависимость от времени, поскольку события случаются в какой-то момент времени, а состояния сохраняются в течение какого-то промежутка времени. Связь фактов и их временных показателей определяется так называемой *темпоральностью*, которая в общем понимании подразумевает временную сущность фактов, определение времени их совершения или динамики их движения и изменения по временной шкале.

С течением времени проверенный факт может быть опровергнут или определен с ограничениями. Если говорить о событиях, то может произойти $F(E) \neq F(E)$, где F – факт; E – событие. С добавлением дополнительного атрибута *время действительности* можно точно определить, в какой момент времени факт был истинным или ложным: $F(E, t_1) \neq F(E, t_2)$, где F – факт; E – событие; t_1 и t_2 – моменты времени.

Введем следующую классификацию фактов.

1. Факты о совершенных событиях.

Это факты, связанные с отдельными событиями, имеющими место быть во временных точках (рис. 3).

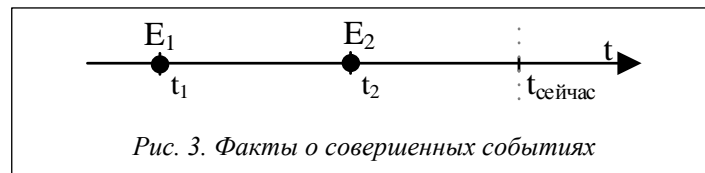


Рис. 3. Факты о совершенных событиях

Примеры: факт о событии списания денежной суммы с карты – факт $F(E_1, t_1)$; факт о событии начисления денежных средств на карту – факт $F(E_2, t_2)$.

2. Факты о сменившихся состояниях.

Это факты, связанные с отражением состояний действительности на временных интервалах (рис. 4).

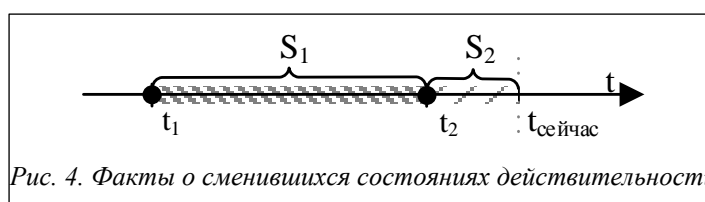


Рис. 4. Факты о сменившихся состояниях действительности

Примеры: факт о состоянии активности банковской карты в течение интервала $[t_1, t_2]$ – факт $F [S_1, t_1, t_2]$; факт о состоянии блокировки банковской карты в течение интервала $[t_2, t_{\text{сейчас}}]$ – факт $F [S_2, t_2, t_{\text{сейчас}}]$.

3. Факты о прогнозируемых событиях.

Это факты, связанные с событиями, которые прогнозируемо совершатся во временных точках в будущем (рис. 5).

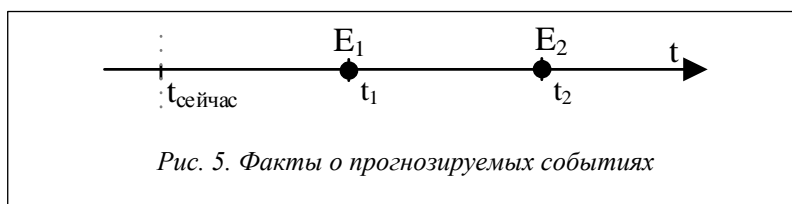


Рис. 5. Факты о прогнозируемых событиях

Примеры: «солнце взойдет завтра утром» – факт $F (E_1, t_1)$; «завтра я пойду на работу» – факт $F (E_2, t_2)$.

4. Факты о прогнозируемых состояниях.

Это факты, связанные с прогнозируемым отражением будущих состояний действительности на временных интервалах (рис. 6).

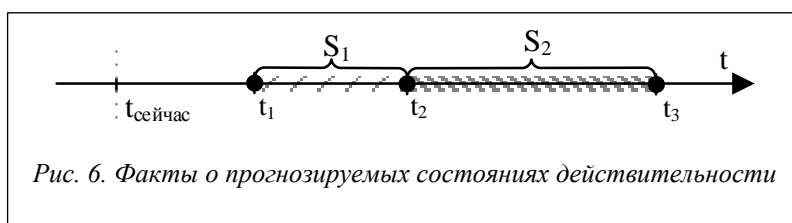


Рис. 6. Факты о прогнозируемых состояниях действительности

Примеры: «солнце будет светить с утра до вечера» – прогнозируемый факт $F [S_1, t_1, t_2]$; «далее будет ночь» – прогнозируемый факт $F [S_2, t_2, t_3]$.

Отношение R_t может называться темпоральным, если в нем существует временной интервал, определяющий период действия (состояния) экземпляра кортежа данного отношения.

Общий вид темпорального отношения R_t :

$$R_t = A \cup I,$$

где A – множество нетемпоральных атрибутов $\{a_1, a_2, \dots, a_n\}$ нетемпорального домена $D \{D_1(a_1), D_2(a_2), \dots, D_n(a_n)\}$; I – множество темпоральных атрибутов $\{t_s, t_e\}$ темпорального домена T , определяющих интервал действия кортежа данного отношения.

Первичным ключом темпорального отношения R_t является комбинация нетемпоральных атрибутов кортежа отношения и интервала его действия:

$$PK(R_t) = A_{pk} \cup I,$$

где A_{pk} является множеством или элементом множества нетемпоральных атрибутов A отношения R_t .

Примером темпорального отношения может быть таблица, в которой два атрибута времени строго определяют интервал действия состояния.

Построение модели хранения и обработки нечетко-темпоральной информации

Рассмотрим одно из разработанных отношений «Перед», представленное в таблице.

Отношения возможных точек и интервалов

Наименование отношения	Точки a, b Интервалы IPeriod, JPeriod	Вид
Перед (Before)	a перед b	PosBefore (a, b[, PExtent_IN])
	IPeriod перед JPeriod	PosBefore (IPeriod, JPeriod[, PExtent_IN])
	a перед IPeriod	PosBefore (a, IPeriod.startdate)

Данное отношение предназначено для определения наступления возможностной временной точки a до возможностной временной точки b .

Для расчетов применяется функция принадлежности теории нечетких множеств, которая позволяет определить степень принадлежности элементов фундаментального множества пространства рассуждения нечеткому множеству. За основу меры степени принадлежности взяты введенные Л. Заде [3] веществен-

ные числа в интервале от 0 до 1. Значение 0 означает, что элемент не включен в нечеткое множество, 1 – описывает полностью включенный элемент. Значения между 0 и 1 характеризуют нечетко включенные элементы.

Фундаментальным множеством является отношение $R = AxB$, то есть подмножество декартового произведения дискретных множеств $A = \{a_1, a_2, \dots, a_n\}$ и $B = \{b_1, b_2, \dots, b_n\}$, где a_n и b_n – возможные временные точки. Нечеткое подмножество R_p строится посредством функции принадлежности $\mu_{R_p}(x)$, которая определяет, в какой степени каждый из элементов фундаментального множества $x \in R$ принадлежит нечеткому подмножеству R_p , то есть принимает следующий вид:

$$R_p = \{(x, \mu_{R_p}(x)) / x \in R\}.$$

Отношение «Перед» принимает следующий вид:

$$PosBefore(a, b[, PExtent_IN]),$$

где a и b – сравниваемые значения возможных временных точек; $PExtent_IN \in [0, 1]$ (опционально) – требуемая для выполнения условия минимально допустимая степень принадлежности элементов множества R нечеткому подмножеству R_p .

Отношение возвращает ($PExtent_OUT$ [, *boolean*]), где $PExtent_OUT \in [0, 1]$ – степень принадлежности элементов множества R нечеткому подмножеству R_p ; *boolean* – признак выполнения: *true*, если $PExtent_OUT \geq PExtent_IN$; *false*, если $PExtent_OUT < PExtent_IN$.

Пример реализации отношения «Перед» для значений возможных временных точек описан далее.

Тип данных точки a : точное событие

Тип данных точки b : Значение в диапазоне не определено

Возможное событие b с возможным диапазоном времени наступления события b_a и b_b и возможной степенью $= b_PExtent_IN$

$$PosBefore = \begin{cases} 0 * b_PExtent_IN, & b_b \leq a \\ \left(1 - \frac{a - b_a}{b_b - b_a}\right) * b_PExtent_IN, & b_a \leq a < b_b \\ 1 * b_PExtent_IN, & a < b_a \end{cases}$$

Выполняемую процедуру с входными датами демонстрирует фрагмент кода реализованной процедуры «Перед»:

```

DECLARE
    result          BOOLEAN;
    i_numresult     NUMBER (3,2);
    a possibledate;
    b possibledate;
    PExtent        NUMBER (3,2) := 0.4;
    extent_result  NUMBER (3,2);
BEGIN
    a:= POSSIBLEDATE('1',to_date ('2015-05-25 10:33:00','yyyy-mm-dd hh24:mi:ss')
    ,null,null,null);
    b:= POSSIBLEDATE('3',null,null,RANGE_FUZ_DEF(to_date ('2015-05-25 10:00:00'
    ,'yyyy-mm-dd hh24:mi:ss'),to_date ('2015-05-25 11:00:00'
    ,'yyyy-mm-dd hh24:mi:ss'),0.41),null);
    i_numresult:= FUZ_UTIL.posbefore(a,b);
    DBMS_OUTPUT.PUT_LINE ('PosBefore extent = '|| i_numresult);
    result:= FUZ_UTIL.POSBEFORE(a,b,PExtent, extent_result);
    DBMS_OUTPUT.PUT_LINE ('PosBefore result = '|| CASE result WHEN true THEN
    'True' ELSE 'False' END ||
    ' where resulted extent = ' || extent_result);
END;
/

```

Результат выполнения отражен на рисунке 7.

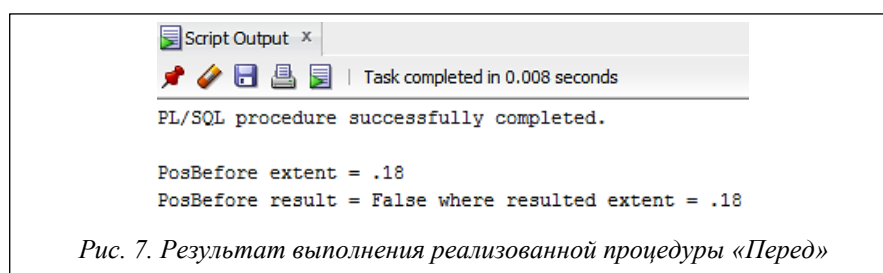
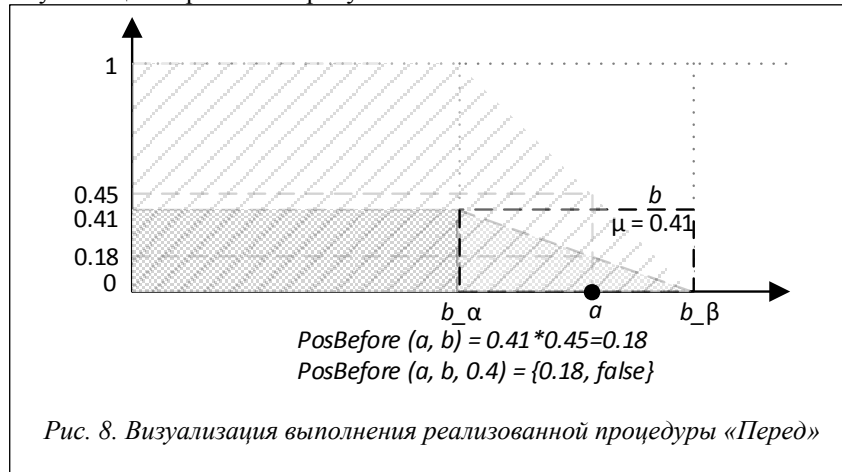


Рис. 7. Результат выполнения реализованной процедуры «Перед»

Графическая визуализация отражена на рисунке 8.



В случаях, когда b_α и/или b_β принимают значение null, для проведения корректных расчетов функция будет подставлять даты начала или завершения значениями по умолчанию (d_{sd} и d_{ed}), которые должны быть определены в соответствии с предметной областью приложения, вызывающего данную функцию.

При $b_\alpha = \text{null}$ и $b_\beta \diamond \text{null}$:

$$PosBefore = \begin{cases} 0 * b_PEntent_IN, & b_\beta \leq a \\ \left(1 - \frac{a - d_{sd}}{b_\beta - d_{sd}}\right) * b_PEntent_IN, & d_{sd} \leq a < b_\beta \\ 1 * b_PEntent_IN, & a < d_{sd} \end{cases}$$

При $b_\alpha \diamond \text{null}$ и $b_\beta = \text{null}$:

$$PosBefore = \begin{cases} 1 * b_PEntent_IN, & a \leq b_\alpha \\ \left(1 - \frac{a - b_\alpha}{d_{ed} - b_\alpha}\right) * b_PEntent_IN, & b_\alpha \leq a < d_{ed} \\ 0 * b_PEntent_IN, & d_{ed} < a \end{cases}$$

При $b_\alpha = \text{null}$ и $b_\beta = \text{null}$:

$$PosBefore = \begin{cases} 0 * b_PEntent_IN, & d_{ed} \leq a \\ \left(1 - \frac{a - d_{sd}}{d_{ed} - d_{sd}}\right) * b_PEntent_IN, & d_{sd} \leq a < d_{ed} \\ 1 * b_PEntent_IN, & a < d_{sd} \end{cases}$$

Библиотека алгоритмов обработки нечеткой темпоральной информации

Разработанная алгебра отношений возможных временных точек и/или интервалов была реализована посредством библиотеки алгоритмов в реляционной БД. Рассмотрим реализацию одного из отношений «Перед». Типы данных, определяющие значения возможных временных точек, следующие:

```
create or replace TYPE "POSSIBLEDATE" AS OBJECT
(
  typeid VARCHAR2(1), -- 1-date type; 2-fuzzy date; 3-range with extent; 4
  eventdate DATE,
  fuzdate DATE_FUZ,
  rangefuzdate RANGE_FUZ_DEF,
  rangenotdate RANGE_NOT_DEF
);
```

К примеру, тип данных «Значение определено с возможным отклонением» (DATE_FUZ) принимает следующий вид:

```
create or replace TYPE "DATE_FUZ" AS OBJECT
(
  ltime VARCHAR2(8), -- 'hh24:mi:ss'
  eventdate date,
  rtime VARCHAR2(8) -- 'hh24:mi:ss'
);
```

Отношение PosBefore было реализовано с использованием средств языка pl/sql. Данное отношение обрабатывает принимаемые на вход темпоральные или возможностьно-темпоральные элементы и возвращает на выходе результат операции в соответствии с заданной ранее спецификацией. Представим фрагмент кода:

```

FUNCTION POSBEFORE(
  a IN possibledate,
  b IN possibledate)
RETURN NUMBER
IS
  v_numresult NUMBER (3,2);
  a_t2_left DATE := a.fuzdate.eventdate - FUZ_UTIL.FUZGETSECONDS(a.fuzdate.ltime)/(24*60*60);
  a_t2_right DATE:= a.fuzdate.eventdate + FUZ_UTIL.FUZGETSECONDS(a.fuzdate.rtime)/(24*60*60);
  b_t2_left DATE := b.fuzdate.eventdate - FUZ_UTIL.FUZGETSECONDS(b.fuzdate.ltime)/(24*60*60);
  b_t2_right DATE:= b.fuzdate.eventdate + FUZ_UTIL.FUZGETSECONDS(b.fuzdate.rtime)/(24*60*60);
  type a_t2_date_array_t is table of date;
  type b_t2_date_array_t is table of date;
  a_t2_date_array a_t2_date_array_t := a_t2_date_array_t (a_t2_left, a.fuzdate.eventdate, a_t2_right);
  b_t2_date_array b_t2_date_array_t := b_t2_date_array_t (b_t2_left, b.fuzdate.eventdate, b_t2_right);
  pos_extent number;
  a_t3_left date := default_start_date;
  a_t3_right date := default_end_date;
  b_t3_left date := default_start_date;
  b_t3_right date := default_end_date;
  type a_t3_date_array_t is table of date;
  type b_t3_date_array_t is table of date;
  a_t3_date_array a_t3_date_array_t;
  b_t3_date_array b_t3_date_array_t;
BEGIN
  --checking values for Type '3', if NULL - then assigning to default values
  IF a.rangefuzdate.startdate is NOT null
  THEN a_t3_left := a.rangefuzdate.startdate;
  END IF;
  IF a.rangefuzdate.enddate is NOT null
  THEN a_t3_right := a.rangefuzdate.enddate;
  END IF;
  a_t3_date_array := a_t3_date_array_t (a_t3_left, a_t3_right);

```

Пусть в таблице fuzdatetypes хранятся данные об объекте с ID '1', который принимает определенное состояние STATE='10' в период с точной даты START_DATE=POSSIBLEDATE('1',2015-05-25 18:54:22.0',null,null,null) по возможностьную дату END_DATE=POSSIBLEDATE('2',null,DATE_FUZ('00:01:00',2015-05-25 18:56:00.0','00:01:00'),null, null).

Используя отношение POSBEFORE, можно определить, находится ли значение START_DATE перед значением END_DATE с возвращаемой степенью возможности. Помимо этого, отношение может проверять принимаемую на вход степень возможности, требующуюся для удовлетворения условия и возвращать соответствующий результат.

Пример выполнения отношения POSBEFORE:

```

SET serveroutput ON;
DECLARE
  result          BOOLEAN;
  i_numresult     NUMBER (3,2);
  startdate      possibledate;
  enddate        possibledate;
  PExtent        NUMBER (3,2) := 0.44;
  extent_result  NUMBER (3,2);
BEGIN
  SELECT ft.start_date INTO startdate FROM fuzdatetypes ft WHERE ft.ID = '1';
  SELECT ft.end_date INTO enddate FROM fuzdatetypes ft WHERE ft.ID = '1';
  i_numresult:= FUZ_UTIL.posbefore(startdate,enddate);
  DBMS_OUTPUT.PUT_LINE ('PosBefore extent = ' || i_numresult);
  result:= FUZ_UTIL.posbefore(startdate,enddate,PExtent,extent_result);
  DBMS_OUTPUT.PUT_LINE ('PosBefore result = ' || CASE result WHEN true THEN 'True' ELSE 'False' END ||
  ' where resulted extent = ' || extent_result);
END;
/

```

Результат выполнения отношения POSBEFORE:

```
PosBefore extent = 1
```

```
PosBefore result = True where resulted extent = 1
```

Апробация алгоритмов

Апробация библиотеки алгоритмов возможностных временных точек и/или интервалов осуществлялась в системах телекоммуникации. Был проведен анализ работы биллинговых систем оператора связи и их взаимодействия со смежными системами. В ходе анализа выявлены процессы, в которых присутствуют неопределенность и неточность в информации, связанной с временными показателями. Оператор связи ежемесячно предоставляет абонентам премиальные бонусные пакеты вида «Скидка 25 % на пользование услугами Интернет». Возможность предоставления премиального пакета рассчитывается специализированной системой лояльности оператора на основании сумм ежемесячных платежей абонентов. Премиальные пакеты предоставляются раз в месяц после получения информации о платежах абонентов из системы биллинга оператора.

Проблема. В соответствии с публичной офертой оператор обязуется предоставлять абонентам пакеты ежемесячно, не позднее обозначенной даты. Существует ограничение, не позволяющее оператору выполнить обязательства: ввиду больших объемов данных предоставление премиальных пакетов происходит с задержками, так как заранее неизвестны точная дата и время завершения расчетного периода в биллинге, по окончании которого информация может быть передана в систему лояльности. Задержки в предоставлении пакетов приводят к негативным отзывам и жалобам от абонентов, тем самым увеличивая показатель оттока.

Задача. Предоставлять абонентам информацию о предполагаемой дате активации премиальных пакетов в соответствии с суммой их платежей.

Решение. Применение библиотеки алгоритмов возможностных временных точек и/или интервалов позволило фиксировать возможностные даты получения информации о платежах каждого абонента, тем самым позволив абонентам получать информацию о возможностной дате предоставления премиального пакета.

Процедура:

– абонент аутентифицируется в онлайн-кабинете и переходит в раздел бонусной программы для проверки причин отсутствия премиальных пакетов;

– онлайн-кабинет взаимодействует с системой лояльности для получения информации о пакетах абонента; ввиду отсутствия предоставленных пакетов в текущем месяце система лояльности обращается в биллинг с запросом информации о платежах абонента, в соответствии с объемами данных и показателями производительности биллинг возвращает возможностную дату завершения расчетного периода;

– система лояльности фиксирует информацию в БД, отображает соответствующие даты в онлайн-кабинете, а также направляет email и SMS-уведомление абоненту, информирующее его о предполагаемой дате предоставления премиального пакета;

– абонент получает информацию и ожидает указанного периода.

Таким образом, в настоящей работе начаты построение и реализация модели обработки темпоральных данных, в которых присутствует неточность или неопределенность, а также подхода к обработке возможностных отношений.

В дальнейшем исследовании планируется продолжить построение новой возможностно-темпоральной модели, а именно: определить и описать все возможные отношения между возможностными временными точками и интервалами, изучить возможностные нетемпоральные периоды и применение нескольких возможностных временных осей в одном отношении, а также реализовать модель для языков определения и манипулирования данными (DML, DDL).

Литература

1. Allen J.F. Maintaining knowledge about temporal intervals. In Communications of the ACM, 1983, pp. 832–843.
2. Nagypál G., Motik B. A fuzzy model for representing uncertain, subjective, and vague temporal knowledge in ontologies. On The Move to Meaningful Internet Systems. 2003, pp. 903–923.
3. Заде Л. Понятие лингвистической переменной и его применение к принятию приближенных решений. М.: Мир, 1976. 165 с.