

ИНФОРМАЦИОННАЯ ПОДДЕРЖКА ПРЕЦЕДЕНТНОГО ПОДХОДА В СИСТЕМАХ ПРИНЯТИЯ РЕШЕНИЙ ПРИ ЧРЕЗВЫЧАЙНЫХ СИТУАЦИЯХ

Олександр Куземин, Михаил Щукин, Богдан Ткаченко, Мария Ткаченко

Abstract: *In this work we described database for emergencies prediction and decision making system, also we designed intellectual knowledge base for real-time data population and intellectual analysis*

Keywords: *emergencies, natural disaster, knowledge base, database, intellectual system, decision making*

ACM Classification Keywords: *H.3.3 Information Search and Retrieval*

Введение

Широко известные проблемы предупреждения и ликвидации последствий чрезвычайных ситуаций (ЧС) определяются в первую очередь недостатком времени для анализа ситуации, а во вторую качеством исходной информации для анализа ситуации и принятия адекватных управляющих решений. Создание ситуационных центров (СЦ) [1] и систем поддержки принятия решений (СППР) [2] требуют значительных усилий для решения следующих проблем:

- анализ методов и моделей при выделении перечня **релевантной** информации для баз данных (БД) и базы знаний (БЗ) с использованием как количественных, так и качественных показателей;
- выявление **при поиске в Internet и социальных сетях** полезной и достоверной информации из текстовой информации (от очевидцев, средств массовой информации, ведомственных отчетов, исторических сведений и т.п.) для наполнения БЗ [3];
- обоснование надежных **средств и форм представления, интерпретации, сбора, обработки, анализа и отображения** разнородной накапливаемой информации в СЦ и СППР.

Постановка задачи для исследования

В общем виде мониторинг ЧС предусматривает получение в результате наблюдений и измерений **данных** (факты, характеризующие отдельные свойства объектов, процессов, явлений), **знаний** (закономерности, связывающие данные и позволяющих решать поставленные задачи в зоне ЧС, полученные очевидцами ЧС и экспертами из практической деятельности) и **контент** (неструктурированные документы, карты, рисунки, схемы, и т.п.).

Рассмотрим задачу исследования для природной ЧС. На начальном этапе сбора информации о ЧС фиксируются основные параметры:

- геоинформационная привязка зоны, в которой возможно возникновение ЧС;
- характеристики причин возникновения ЧС, тип/вид и площадь ЧС;
- данные мониторинга окружающей среды, которые снимаются и вносятся для анализа, обработки и хранения в БД;
- оценки текущего социально-экономического ущерба, количество пострадавших людей, характеристики разрушений и оценки прогноза возможного развития ЧС;
- характеристика сценария развития ЧС, оценки риска для жизнедеятельности людей и расчетные параметры плана уменьшения ущерба, количественные оценки объемов необходимых ресурсов, сил и средств, которые необходимо для локализации и ликвидации последствий ЧС;
- сравнительные оценки возможности использовать прецедентный опыт.

Этот далеко не полный перечень оценочных и фактических данных и знаний, которыми надо обеспечить СППР для предупреждения и развития, ликвидации последствий ЧС. Для решения этих проблем необходима разработка информационной модели ЧС, которая позволит с минимальными затратами времени и средств учесть динамику развития ЧС и существующий прецедентный опыт. Важными аспектами предлагаемого подхода являются:

- использование для принятия решений в СППР расширенной БЗ о прошедших ЧС из Internet и социальных сетей [3];
- применение рассуждений по аналогии (по прецедентам, case-based reasoning) [4], данный подход является реализацией метода интеллектуального анализа данных (Data Mining).

Использование Internet и социальных сетей, прежде всего, позволяют получить информацию тогда, когда в месте наблюдения до и после ЧС можно использовать мнение значительного числа экспертов. Кроме того, свидетельства могут дать информацию в том случае, когда нет централизованных средств получения информации.

Конечно, есть проблема, которая связана с тем, что не всегда можно доверять качеству знаний полученных таким способом. Предлагаемые методы формализации, моделирования и анализа позволяют с определенной степенью достоверности получить значительное число качественных знаний. Исследуемая модель прецедента обладает свойством адаптивности к условиям поиска решений на основе модели предметной области и обобщенной структуры прецедента, которая

позволяет решить многокритериальную задачу оптимизации управленческих решений в условиях неопределенности [4].

Технология разработки моделей прецедента имеет следующие шаги:

Шаг 1. Строится модель предметной области для контролируемой зоны ЧС.

Шаг 2. Тестируется модель для заданных критериев при решении различных задач анализа и принятия решений в зоне ЧС. Решаются тестовые задачи многокритериальной оптимизации в условиях неопределенности.

Шаг 3. По результатам оптимизации корректируется модель предметной области (расширяется или дополняется).

Шаг 4. Разрабатываются модели прецедента для контроля, анализа, предупреждения, снижения ущерба и ликвидации последствий ЧС.

Шаг 5. Отбор в прецедентную БЗ наиболее удачных решений, которые с наилучшими показателями снижения риска безопасности жизнедеятельности людей, социального риска, материальных потерь и т.п.

Шаг 6. Оценивается близость контролируемой ЧС к прецеденту.

Шаг 7. Адаптируются решения по заданным критериям.

Шаг 8. В случае не удовлетворительного результата при поиске решений в БД и БЗ ищется новый прецедент, и процесс решения возвращается к шагу 5.

Для решения поставленной задачи при разработке информационной модели ЧС используем понятия микроситуации [5] и вывод на основе прецедентов [6].

Технология системного анализа и выбора управленческих решений с использованием прецедентов для предупреждения, ликвидации последствий ЧС, снижения затрат времени, используемых средств, уменьшения социальных рисков и потерь при возникновении ЧС в контролируемой зоне ЧС предлагается выполнять в следующей последовательности:

Шаг 1. Выполняется временной и дисперсионный анализ влияния климатических параметров $\{X\}$ окружающей среды в зоне ЧС. Формируется множество микроситуаций – $\{Msit\}$ для контролируемой ЧС.

Шаг 2. Ранжируются микроситуации – $Msit$ из множества микроситуаций – $\{Msit\}$ по степени влияния на возникновение ЧС [6].

Шаг 3. Из данных о прошедших ЧС формируется БД микроситуаций для контролируемой ЧС

$$\left\{ \text{Msit} \left\langle \text{Geo}, X, \mathfrak{S}, \text{Res}, R_{\text{ef}} \right\rangle \right\},$$

где Geo – параметры геоинформационной привязки;

\mathfrak{S} – необходимые управленческие решения;

Res – требуемые ресурсы в условиях ЧС;

R_{ef} – эффект от снижения риска

$$R_{\text{ef}} = \frac{R_{\text{before}} - R_{\text{after}}}{C_{\text{var}}},$$

где R_{after} – подверженность риску до возникновения ЧС и после возможного применения требуемого управленческого решения и снижения риска; C_{var} – затраты, связанные с применением данного управленческого решения (причем управленческое решение приемлемо при значении $R_{\text{ef}} > 1$).

Шаг 4. Поиск для контролируемой ЧС формируются текущие микроситуации и осуществляется поиск близких для $\left\{ \text{Msit} \left\langle \text{Geo}, X, \mathfrak{S}, \text{Res}, R_{\text{ef}} \right\rangle \right\}$ контролируемой ЧС аналогов из базы прецедентов $\left\{ \text{Msit}^* \left\langle \text{Geo}^*, X^*, \mathfrak{S}^*, \text{Res}^*, R_{\text{ef}}^* \right\rangle \right\}$. Сравнение осуществляется на основе расстояния, вычисляемого по метрике Хемминга. Близость прецедентов-аналогов ЧС определяется как произведение близости всех микроситуаций.

Шаг 5. Для микроситуаций, которые расположены по шкале ранжирования в верхней позиции шкалы (т.е. имеют наибольшее влияние на возникновение ЧС) определяются наилучшее управляющее решение для управления рисками – \mathfrak{S}^* и выбранные при этом эффективные ресурсы Res^* .

Шаг 6. В экспертной системе СППР оцениваются, рекомендуемые по результатам предыдущего шага управляющие решения.

Шаг 7. Результаты передаются лицу принимающему решение (ЛПР) в СЦ.

Упомянутая выше проблема расширения информации о ЧС за счет поиска ее в Интернет и социальных сетях вызывает необходимость раскрытия технологии наполнения БЗ. Необходимо решить проблемы организации поиска информации и использовании этой информации при моделировании предметной области.

Необходимо отметить, что знания, выявленные с помощью Интернет и социальных сетей, имеют вербальное описание, которое может обладать детализированным перечнем возможных отклонений от нормы смысловой определенности:

1. Приблизительное представление о некоторых процессах при ЧС, состоянии объектов, событиях, понятиях, отношениях и связях и т.п.
2. Двусмысленность, допускающая двоякое истолкование.
3. Неточность — нет подробного указания признаков.
4. Слабая детализированность.
5. Формулировки даны в общих терминах, охватывающих довольно много различных случаев, со значительными погрешностями или вообще не правильных.
6. Описание не очень полное, нет ответа на некоторые, сформулированные выше вопросы.

Кроме того, развитие технологии поиска решений за счет использования понятия микроситуации [5].

Решению некоторых значимых проблем, которые упоминались выше и посвящается данное исследование.

Особенности разработки базы знаний с использованием интернет

В последнее время в предотвращении и ликвидации последствий ЧС стали играть принципиально новые субъекты гражданского общества в виде локальных структур – комьюнити (community) или сетевых сообществ. Это связано с бурным развитием социальных сетей и необходимостью своевременного выявления и получения дополнительных сведений о ЧС, которые не могут быть реализованы в существующих системах управления в условиях ЧС. Социальные ресурсы сетевых сообществ позволяют им реализовать ряд функций, с помощью которых происходит актуализация сетевых сообществ как субъектов действий в условиях ЧС [9]. Разработка БЗ будет формироваться и анализироваться в экспертной системе СППР с учетом ввода текстовой информации из Интернет и социальных сетей.

В результате проведения объектно-структурного анализа зоны ЧС для заполнения БЗ необходимо получить ответы о предметной области, связанной с ЧС на следующие вопросы:

ЗАЧЕМ нужны знания (полученные в результате стратегического анализа ЧС: представление о типе ЧС, функциях субъектов и объектов в зоне ЧС);

КОМУ и **КТО** готовит знания (полученные в результате организационного анализа: для службы мониторинга, спасательных организаций, структуры министерства по чрезвычайным обстоятельствам и т.п.);

ЧТО представляют собой знания (полученные в результате концептуального анализа: концепты и отношения);

КАК представлены знания (в результате функционального анализа: гипотезы, правила вывода, факты и модели для СППР);

ГДЕ происходит ЧС – знания (поступившие в результате пространственного анализа: окружающая ЧС среда, информация об объектах и оборудовании в зоне ЧС);

КОГДА поступают знания (в результате временного анализа: временные параметры событий сопровождающих ЧС, критические данные и ограничения);

ПОЧЕМУ произошла ЧС – знания (поступившие в результате каузального анализа для формирования подсистемы объяснений в СППР);

СКОЛЬКО ожидается или требуется затрат, ожидается потерь от происшедшей ЧС знания (поступившие в результате экономического анализа о том какие экономические, социальные и материальные ресурсы необходимы для смягчения и ликвидации последствий ЧС и требуемые коммуникации).

Ответы на сформулированные выше вопросы могут быть получены в результате опроса очевидцев ЧС. Однако, для качественного анализа при формировании БЗ требуют специальных методов обработки и моделей представления. Наиболее предпочтительными могут быть опросные, структурированные последовательности вопросов. В случае не возможности провести такой опрос необходим лингвистический и семантический анализ опубликованных текстовых сообщений.

Сбор информации в прецедентную БЗ либо на основе диалога с очевидцами (задаются в определенной последовательности вопросы типа «какой тип ЧС?», «что Вы наблюдаете?», «где?», «когда...?», «сколько...?» и т.п. Используются различные типы шаблонные ответы, из которых «очевидец» может выбрать ответ.

Остановимся на особенностях выявления знаний при общении с компьютером. В этом случае интеллектуально-вербальная коммуникация с очевидцами ЧС основана на связи объекта наблюдения – ЧС, поведения – представления динамики в предметной области и слова –

описания наблюдаемых событий. В данном случае некоторый физический или абстрактный объект из окружающей среды является некоторым понятием и выражается определенным словом. Данная связь схематически представлена на (Рис 1.)



Рисунок 1 – Связь объекта, понятия и слова

Процессы словесного озвучивания объектов и восприятия слов изображены на (Рис. 2).

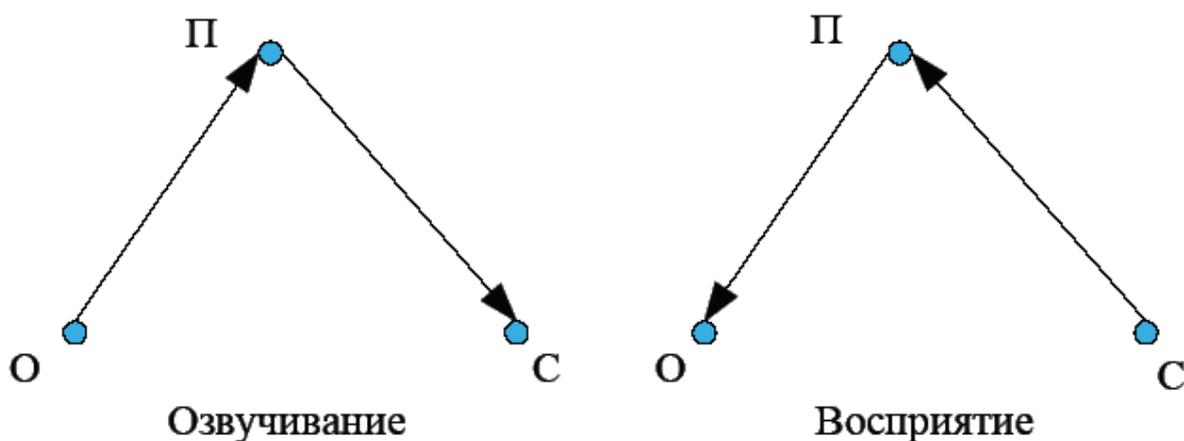


Рисунок 2 – Озвучивание объектов и восприятие слов

Человек представляя определенный объект в своем сознании находит соответствующее ему понятие, к которому затем подбирает подходящее слово и озвучивает его. Восприятие текста происходит в обратном направлении. Получив через органы коммуникации некоторое слово, человек связывает его с определенным понятием в зависимости от контекста ситуации, своего опыта и других факторов. После формирования в сознании определенного понятия формируется представление о ЧС.

Проблемы в интеллектуально-вербальной коммуникации возникают при использовании неверных слов. Это происходит в случае наличия различных ассоциаций между понятием и словом. Один участник коммуникативного процесса выражает понятие одним словом, с которым у другой стороны ассоциируется другое понятие. На (Рис. 3) отображена проблема, которая характеризует обмен информацией между двумя лицами 1 и 2. В данном случае имеются две связки $O-P-C-O_1-P_1-C_1$ и $O_2-P_2-C_2$ у каждого из участников соответственно. Первый из них считает, что слово C_2 выражает понятие P_1 . Второй, будучи уверенным, что понятию P_1 соответствует слово C_1 , воспринимает слово C_2 как выражение понятия P_2 . Поэтому в данном информационном обмене наблюдается неверная передача объекта O_1 , который принимающая сторона воспринимает как объект O_2 .

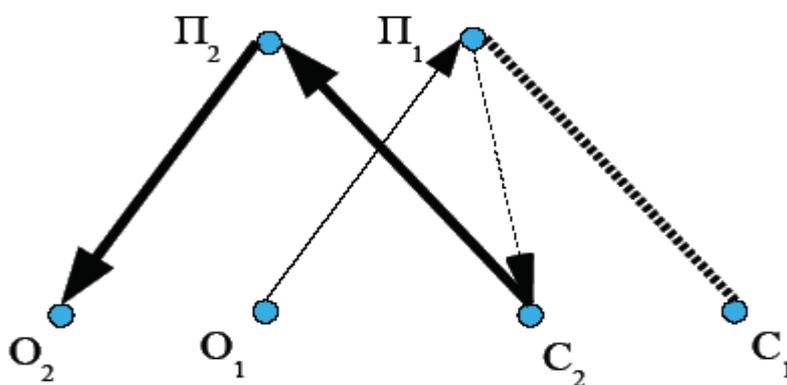


Рисунок 3. – Ошибка в выборе слова

Описанная проблема особенно актуальна на начальных стадиях заполнения БЗ, когда происходит описание окружающей среды на основе поиска знаний в Интернет. Необходимо уточнить задачи, которые стоят перед инженером по знаниям в экспертной системе (ЭС) СППР. Разногласия на данном этапе ведут к возникновению серьезных ошибок.

Однако не смотря на вышеописанные проблемы текстовое представление является одним из наиболее эффективных с выразительной точки зрения и с точки зрения затрат ресурсов на подготовку и создание описания. Для $C_2 O_2 C_1 O_1 P_1 P_2$ можно сформулировать две гипотезы относительно представления информации об объектах связанных с ЧС.

Гипотеза 1. Всю информацию об объекте управления, ресурсах, используемых при смягчении влияния и ликвидации последствий ЧС можно выразить средствами обычного естественного языка.

Гипотеза 2. Всякий текст, описывающий объект управления может быть переведен на формальный язык.

Безусловно, для некоторых случаев при возникновении ЧС могут существовать объекты, для которых гипотезы не верны, однако опыт показывает справедливость приведенных гипотез и целесообразность их использования.

ЧС состоит из множества понятий ее образующих. Как было отмечено выше, для поиска эффективного управленческого решения в условиях контролируемой или наблюдаемой ЧС необходимо провести сравнение данной контролируемой ситуации с множеством прецедентных из БД и БЗ.

На начальном этапе моделирования предметной области формируется идентификация понятий, входящих в нее. Классические методы классификации, такие как кластерный анализ, в чистом виде не подходят, потому что при моделировании предметной области заранее неизвестно, что собой представляют понятия, которые используются в модели ЧС и как их атрибуты соотносятся с атрибутами прецедентных понятий. Поэтому перед описанием модели ситуации требуется разработка и проведение ряда дополнительных исследований по идентификации понятий.

В первом приближении наиболее подходящие для сравнения ситуаций и понятий можно сделать на коэффициенты Жакара, Роджерса и Танимото, но они не учитывают асимметричность подобия и поэтому требуют изменения.

Основная проблема с идентификацией понятий и ситуаций, которая присутствует в контексте рассматриваемой проблемы – это то, что сравниваемые объекты (в зоне контролируемой ЧС и из прецедентной БЗ) могут быть разнородными и имеющих разную размерность. В этом случае неприменимы известные меры сходства без соответствующих доработок. Здесь находят применение выше упомянутый инструмент моделирования предметной области с использованием понятия микроситуация [4–6]. Реально можно определить близость значений

для текущей контролируемой зоны ЧС и накопленных данных и знаний прошедших наблюдений, представленных в виде прецедентной БЗ, т.к. поиск близких значений обретает реальные контуры за счет снижения неопределенности и использования микроситуаций, значимых для поиска решений.

При моделировании предметной области, кроме введенных понятий, необходимо рассматривать процедуры, операции, действия, кроме того необходимо использовать и другие термины. Так большой комплекс действий, приводящий к выполнению, который связан с ликвидацией последствий ЧС, называют «направлением работ». Для привязки действий ко времени употребляются термины «стадия» и «этап». Однако главным системным понятием здесь может оставаться операция. Разложение процесса решения в условиях ЧС только на необходимые направления работ или стадии и этапы выполнения, например, спасательных работ практически всегда связано с неопределенностью и многокритериальностью решаемой проблемы.

В общем случае любое действие в условиях ЧС имеет три основные характеристики:

1. цель действия;
2. описание действия;
3. способ его выполнения.

Все эти характеристики можно представить в виде вопросов, которые способствуют организации действия:

Цель (назначение) – это ответ на вопросы в зависимости от ситуации:

«Зачем?» или «Каков должен быть результат?».

Описание действия – это ответ на вопрос «Что делать?»

Способ выполнения (умение и возможность) – это ответ на вопрос «Как делать?»

Или коротко система вопросов будет выглядеть так: «Что?», «Зачем?», «Как?»

Продумывание решения задачи – это способ организации вопросов, т.е. выбор схемы из достаточно небольшого множества (рис. 4).

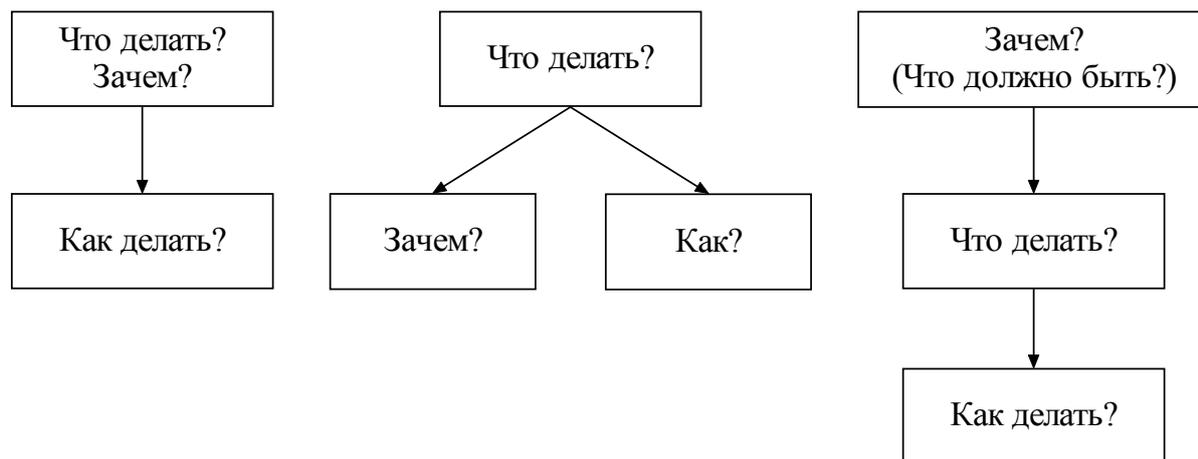


Рисунок 4 – Типовые приемы организации отдельных действий

Однако существуют и ситуации, которые в некотором смысле являются крайними. Их наличие только лишь подчеркивает, что только наличие ответов на все три вопроса приведет к уверенности в правильной организации действий.

Первый случай – это наличие только ответа на вопрос «Как?». Это случай, когда следует действовать **по инструкции в условиях ЧС**.

Второй случай – это пренебрежение ответом на вопрос «Зачем?». Это, как правило, действия по выбору, когда проще перебрать все случаи, чем искать ответ на вопрос «Зачем?». Например: мы что-то пробуем, переделываем и так и эдак, пока не найдем то, что ищем, например, причину потери связи в разрушенном здании или наилучшее решение какой-то проблемы ликвидации последствий ЧС.

Третий случай – это случай пренебрежения вопросом «Как?». Такие действия являются абстрактной схемой, неконструктивным подходом или схоластическим решением. К действиям без «Как?» относятся действия, для которых пока что нет технологии, нет информационных средств с требуемыми ресурсами. Здесь либо требуется специальные исследования, направленные на поиск ответа, на вопрос «Как?», либо выполняется поиск осуществляемого набора действий, заменяющего данный, либо происходит изменение (упрощение) задачи.

Изобразим теперь наиболее распространенную схему организации действий по решению задачи (Рис. 5).

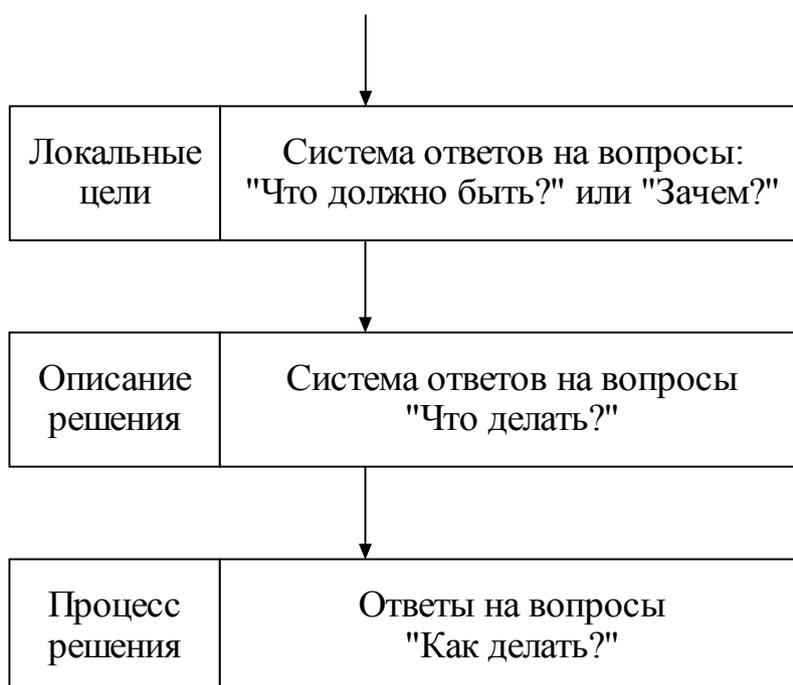


Рисунок 5 – Типичный вариант организации систем действий по решению задачи ликвидации последствий ЧС

Заметим, что нижняя ячейка не является системой (в отличие от двух верхних), так как ответы на вопросы «Как?» не обладают внутренними связями. Замена одного способа выполнения другим обычно не приводит к каким-либо изменениям в решении задачи; однако замена ответа на вопрос «Что?» или «Зачем?» влечет изменение других действий или целей. Тут уместно было бы использовать модели, которые используют Байесовский подход для решения проблемы. Результаты работы, которые представлены в следующей статье, должны расширить возможности решения поставленной проблемы.

Выделим совокупность локальных целей $\{g_j^s\}$, обеспечивающих выполнение цели общей по ликвидации последствий ЧС – G_j с помощью операции J :

$$\{g_j^s\} \rightarrow G_j, s=\overline{1, \sigma}.$$

Данная запись означает лишь достаточность достижения целей множества целей $\{g_j^s\}$ для осуществления общей глобальной цели – G_j ; может существовать и другая совокупность

локальных целей, которая обеспечит достижение той же цели – G_j . Основным смыслом этой записи в том, что можно «забыть» про цель G_j и выполнить более простые цели $g_j^1, g_j^2, \dots, g_j^\sigma$.

Систему локальных целей принято создавать сверху, с введения набора целей первого иерархического уровня. Декомпозиция целей должна сопровождаться их согласованием, которое является главным условием того, что все они вместе приведут к достижению глобальной цели задачи ликвидации последствий ЧС.

В общем случае, который обусловлен сложностью рассматриваемой задачи и структурой связей между локальными целями имеет произвольный характер. В качестве крайних ситуаций можно считать следующие случаи:

- случай, когда выполнение любой цели связано с выполнением каждой из остальных целей достижения одного условия (что соответствует отсутствию иерархических связей между целями);
- случай полной независимости целей, т.е. каждая из них достигается самостоятельно.

Можно выделить простейшие и одновременно основные типы связей между целями – их существует всего три.

Пусть имеется цель G_j , которую можно разделить на две цели g^1 и g^2 (индекс j здесь для простоты опустим). Тогда это можно записать так:

$$\{g^1, g^2\} \rightarrow G_j.$$

При этом возможно:

- **последовательное** выполнение – только достижение одной цели, которое дает возможность выполнять действия для достижения другой цели;
- **параллельное** выполнение – цели могут достигаться независимо;
- **циклическое** выполнение – частичное выполнение одной из множества целей, при этом данное действие позволяет частично достигать другую цель, что в свою очередь позволяет вернуться к выполнению первой цели, и так до выполнения обеих целей.

Цели на рисунке 6.а называются **связанными**, на рисунке 6.б – **несвязанными**, на рисунке 6.с – **сложно-связанными**. Последние с точки зрения кибернетики являются примером системы с обратной связью.

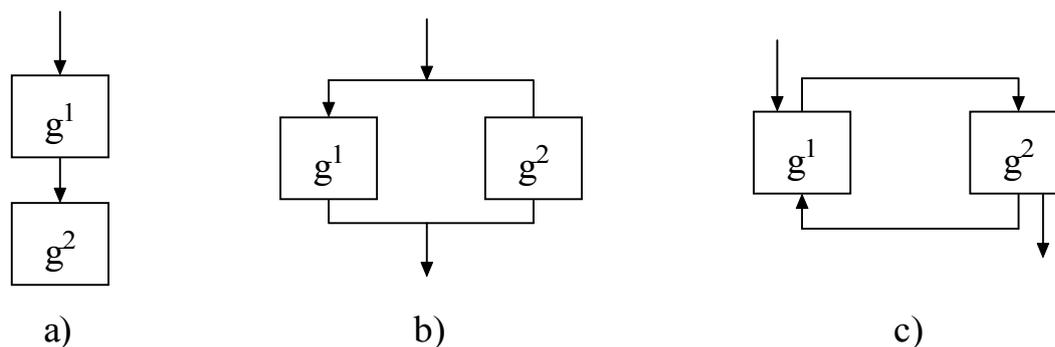


Рисунок 6 – Способы достижения двух целей

Способ достижения каждой из целей g^1 и g^2 в отдельности может быть дискретным (порциями, скачками) и непрерывным. В первом случае схему (Рис. 4.а) также называют интерактивной, а каждый переход от цели g^1 к g^2 и обратно – итерацией, шагом, циклом.

Для более двух локальных целей связь между ними будет комбинированием приведенных выше типов. Схематические примеры некоторых из них для случая трех локальных целей изображены на (Рис. 7).

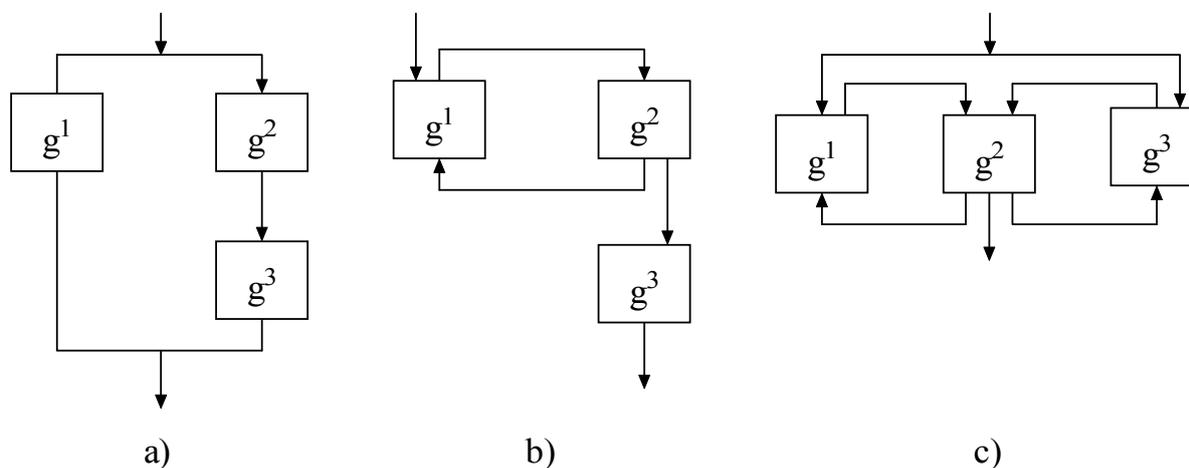


Рисунок 7 – Некоторые способы достижения трех целей

Для более двух локальных целей связь между ними будет комбинированием приведенных выше типов.

Часто выполнение одной локальной цели может затруднить и даже исключить выполнение другой. Такие цели (две и более) называют антагонистическими. В сложных системах практически не удается избавиться от той или иной степени антагонистичности локальных целей.

После анализа текста, который был получен из Интернет или социальных сетей необходимо для правильной идентификации понятий входящих в описание, которые связаны с ЧС необходимо использовать алгоритм категорийной классификации понятий. Этот алгоритм позволяет эффективно соотнести неизвестное понятие к уже известному классу. Данный алгоритм подразумевает наличие групп категорий классов понятий. Эти категории строятся в виде иерархий, которые образованы по родовому признаку.

Выполняя моделирование ЧС, проектировщик строит модель предметной области, которая есть по своей сути отражение в сознании наблюдателя свойств объектов и событий, их отношений в решении задачи исследования ЧС.

С учетом вышесказанного модель ЧС – Mod , как отражение целей ликвидации последствий ЧС, для определения эффективных средств управления при ликвидации последствий ЧС, можно определить в следующем виде [4]:

$$Mod = \langle A, O, G, P, \Theta \rangle,$$

где A - элементы (компоненты, части модели ЧС) системы управления рисками для ЧС – Mod ,

O - отношения (связи) элементов модели,

G - цели подсистем модели,

P - инженер по знаниям и аналитик ЭС (разработчики модели ЧС – Mod),

Θ - язык моделирования (выбранный метод семантического и логического моделирования ЧС).

Задача моделирования ЧС – это удовлетворение целей G при ликвидации последствий ЧС. На начальных стадиях анализа ЧС и принятия решений в ЭС и СППР инженер по знаниям и аналитик ЭС P с помощью различных, в том числе и языковых средств моделирования – Mod проводит исследование Mod , чтобы выяснить, как соотносятся понятия, элементы модели – A которые имеют отношения O для достижения целей системы G .

Во многом успешность данного процесса моделирования ЧС определяется способностью инструментария Mod наиболее полно и достоверно учесть цели системы.

Модели ЧС можно рассматривать как модель, которая состоит из множества понятий, находящихся в отношениях или связях друг с другом, образующих целостность или органическое единство.

Ситуация формируется как множество количественных (из БД) и качественных (из БЗ) микроситуаций. Ситуация Sit представляется как совокупность из n микроситуаций $\{Msit_i\}, i=\overline{1,n}$ [5]:

$$Sit = \{Msit_i\}, i=\overline{1,n}, Msit_i = \langle e_i, O_{e_i} \rangle.$$

Часть модели Mod, которая определяется знаниями из БЗ и может быть представлена парой $Msit^e = \langle e, O_e \rangle$, называется микроситуацией понятия e .

Понятие e называется центральным понятием микроситуации $Msit^e = \langle e, O_e \rangle$, множество $\{O_e\}$ является контекстом центрального понятия e . Центральным является понятие, являющееся центральной смысловой единицей, вокруг которой базируется анализируемая микроситуация. Множество $O_e = \{o_j\}, j=\overline{1,m}$ состоит из понятий, выражающих отношение центрального понятия e с остальными второстепенными понятиями, участвующими в микроситуации для рассматриваемой ЧС.

Ситуация включает множеством понятий, каждое понятие отражает свойства ситуации в одной из множества характеристических категорий [6].

Структура категории представляет собой иерархию понятий, построенную по родовому принципу, правила формирования, которой будут представлены ниже. Вместе с категориями, представляющими сущностные понятия, будут присутствовать категории с понятиями, представляющими отношения.

Категория Kat представляется в виде иерархии с помощью ориентированного графа $Kat = \langle E, \Theta \rangle$ понятий, построенного по родовому признаку, где E – множество вершин графа – множество понятий $\{e_1, e_2, \dots, e_{n-1}\}$ и $A = \{\Theta_1, \Theta_2, \dots, \Theta_{n-1}\}$ – множество дуг графа,

выражающих отношение наследования между двумя понятиями в направлении от наследника к предку. Понятия e_1 является изначальным понятием категории Kat , которое определяет самый общий элемент в иерархии.

Каждое из понятий включает в себя указатель на решающее правило, которое определяет группу признаков, позволяющих отличать данное понятие от других. Признаком понятия является то, что так или иначе, его характеризует, то, в чем они похожи или отличаются друг от друга, как в плане наличия (положительный признак), так и в плане отсутствия (отрицательный признак) [7].

Для нахождения подходящего проектного решения аналитику необходимо выразить *проблемную ситуацию* в терминах понятий существующих категорий. Объекты понятий, описывающие текущую проблемную ситуацию, классифицируются по имеющимся категориям, описываются связи между ними и затем производится поиск прецедентной ситуации в БЗ, близкой к исследуемой. Прецедентные ситуации, с которыми производится сравнение, называются *эталонными ситуациями*.

Этапы поиска эффективных управленческих решений для ликвидации последствий ЧС (с использованием прецедентной БЗ) выполняются в следующей последовательности:

Шаг 1. Из Интернет и социальных сетей в результате работы поискового робота получают тексты наблюдений очевидцев ЧС для текущей ситуации (предметной области) в виде повествовательного текста.

Шаг 2. Выявление из полученного описания множество понятий из имеющихся категорий.

Шаг 3. Построение описания на языке представления ситуаций (моделирование ЧС).

Шаг 4. Поиск в пределах заданного интервала близости эталонной ситуации из прецедентной БЗ, близкой по структуре к проблемной или контролируемой ЧС.

Шаг 5. Сужение поиска путем включения в поиск учета атрибутов объектов понятий.

Сначала текущая ситуация описывается в виде простого повествовательного текста, отражающего по группам категорий участников и среду. Полученный таким образом текст, является основой для построения описания на языке представления ситуаций. Для перехода к нему необходимо, чтобы описание велось на основе понятий из существующих категорий. Это осуществляется посредством классификации понятий, когда понятия, описанные на модельном языке предметной области, замещаются известными понятиями.

Обозначим текстовое описание ситуации, полученное от эксперта или очевидца ЧС как $TxtExp$. Понятие предметной области из первичного текстового представления $etExp$, а понятие, выраженное в терминах имеющихся категорий – e . Процесс выявления понятий $etExp$ из текста $etExp$ является функцией, отображающей текст на множество понятий:

$$EF(TxtExp) \rightarrow \{etExp_i\}$$

где EF – функция выявления понятия, $etExp_i$ – выявленные понятия имеют значение $i=\overline{1,n}$.

При поиске кандидатов в понятия можно руководствоваться классическим методом, т.е. рассматривать стандартные категории объектов и пытаться выделить из полученного описания понятия, которые, по мнению эксперта или аналитика могли бы принадлежать данным категориям [8]. Для этого можно воспользоваться представленными ранее описаниями групп категорий и самих категорий. Следует отметить, что на данном этапе не происходит классификации понятий по категориям. Используется лишь описание групп и категорий, которые могут послужить определяющими для выделения дополнительных понятий из описания ситуации.

Таким образом, анализируя ситуацию – Sit получаем два множества кандидатов: cE - множество понятий, являющихся сущностями и cO - множество понятий, выражающих отношение между другими понятиями.

Первым шагом является выделение из множества кандидатов cE множества центральных понятий – $cKat$. В первую очередь это существительные, который являются подлежащими в предложениях, из которых они были получены. Затем поиск центральных понятий проводится среди остальных элементов множества кандидатов cE . Для выделения центральных понятий следует использовать следующие рекомендации:

- центральное понятие является обобщающим понятием его составных частей (других понятий);
- центральное понятие каким-либо образом управляет деятельностью или состоянием других понятий.

Обособленное центральное понятие может не иметь связей с другими понятиями предметной области, но являться неотъемлемой частью описания ситуации. В данном случае при изъятии этого понятия описание теряет полноту.

Следующим шагом будет выделение связей полученных кандидатов в центральные понятия. Для этого сначала из множества кандидатов cE удаляются элементы множества центральных понятий cE , в результате чего получится множество отношений и второстепенных понятий $cA=cE/cKat$. Задача стоит в выделении подмножества отношений (ассоциаций) $cO \subset c\Theta$.

Ассоциация центрального понятия с другими понятиями может выражать следующие типы активных связей между этими понятиями:

- подчинение – руководство;
- управление – зависимость;
- слежение – уведомление;
- связь – прием и передача других понятий;
- отношение агрегирования – «целое-часть».

Первыми кандидатами на отношения являются выделенные ранее глагольные формы. Для каждого из полученных центральных понятий множества $cKat$ проводится поиск связанных с ним глагольных форм по приведенным выше типам связей. Затем происходит поиск связей, не выраженных явно глагольной формой, но подразумевающихся в контексте найденного Интернет или социальных сетях описания ЧС, приведенного рассказа, описывающего проблемную ситуацию. Это могут быть следующие дополнительные виды ассоциаций:

- следование правилам, указаниям и т.д.;
- описание среды центрального понятия.

Таким образом, руководствуясь списком типов активных и дополнительных связей, получим элементы множества отношений cO . Каждый из элементов множества отношений $o_i \in cO$ будет связан с определенным центральным понятием $c_i \in cKat$, формируя множество микроситуаций $Msit_i = \langle c_i, O_{c_i} \rangle$, в котором элементы множеств O_{c_i} являются элементами множества cO . На данном этапе полученные микроситуации еще не являются полными, поскольку элементам множества cO еще не поставлены в соответствие второстепенные понятия.

В качестве второстепенных понятий могут выступать любые элементы множества кандидатов cE , независимо от того попали ли они в множество центральных понятий $cKat$, множество

отношений сО или ни в одно из них. В большинстве случаев второстепенные понятия получаются из дополнений к глагольным формам, выделенным ранее. В общем случае второстепенные понятия – это те понятия, на которые ссылаются ассоциации. В качестве примера для приведенных выше видов отношений второстепенными понятиями могут такие понятия:

- подчиняющееся или руководящее понятие – работник или начальник;
- управляющее понятие – министерство чрезвычайных ситуаций (МЧС), подчиняющееся – региональная служба МЧС;
- следящее понятие – пост наблюдения МЧС;
- уведомляющее понятие – датчик температуры, свидетельство очевидца;
- прием и передача – почтовый ящик, модем;
- агрегирование – инструменты для спасательных работ и транспортное средство;
- следование правилам – закон, инструкция по выполнению спасательных работ;
- среда – влажность, сила тока, давление и т.п.

Указания на второстепенные понятия у отношений представляются в виде ссылок на данные понятия.

Полученное представление проблемной ситуации в терминах проблемной области необходимо теперь описать с помощью понятий, представленных в имеющихся категориях. Каждый из классов понятий, расположенный в категории, имеет свои отличительные черты и характеристики, по которым данное понятие можно отличить от других. Каждой категории ставится в соответствие *стратегия* выявления понятий. Стратегия представляется в виде ориентированного графа. Вершинам графа, именуемым *точками*, соответствуют классы понятий из категории, которой соответствует данная стратегия, а дугам, именуемым *переходами*, соответствуют различные решающие правила, приводящие к переходу от одной точки к другой.

Решающим правилом является логическое выражение, состоящее из группы вопросов, представляющих собой характеристические признаки понятия. Истинность данного выражения говорит о том, что понятие, к которому были применены данные признаки и принадлежат к данному классу понятий. При добавлении класса понятия в категорию, в решающее правило включается только те признаки, которых нет у родительских объектов, т.е. выделяются признаки,

отличающие его от непосредственного родителя. Решающее правило представляется в следующем виде:

$$DR = \hat{F}_i$$

где F_i - отдельный признак.

При подсчете значения решающего правила DR каждому из признаков F_i ставится в соответствие значение истины, если данный признак присутствует у понятия и ложь в противном случае.

Стратегия идентификации понятий в категории не является отдельным множеством, а всего лишь смысловым объединением над иерархией понятий в категории и их решающими правилами. Выделение такого отдельного представления позволяет отдельно рассматривать стратегию в алгоритмах как самостоятельный элемент. Стратегия не является статическим множеством, она выстраивается в зависимости от текущей структуры категории и понятий ее наполняющих.

Стратегия для категории $Kat = \langle E, \Theta \rangle$ представляется в виде следующего множества:

$$Sit_{Kat} = \langle P, D, p_0, D_0 \rangle,$$

которое состоит из множества точек $P = \{p_1, p_2, \dots, p_n\}$,

множества переходов $D = \{D_i\}$, $i = \overline{1, n}$ в каждую из точек $p_i \in P$,

ссылки на изначальную точку $p_0 \in P$, с которой начинается идентификация,

ссылки на изначальный переход $D_0 \in D$, с помощью которого принимается решение о принадлежности понятия к данной категории.

$D_k = \{d_{j,k}\}$, параметры $j = \overline{1, n}$, $k = \overline{1, n}$ означают, что переход направлен от точки p_j к точке p_k или по-другому от класса понятия e_j к классу понятия e_k .

Данному переходу D_k от некоторого класса понятия к классу понятия e_k соответствует решающее правило DR_k .

Таким образом, видно, что возможность перехода к некоторой точке определяется лишь связанным с соответствующим классом решающим правилом. Направление переходов $D_k = \{d_{jk}\}$ обратно направлению дуг из множества $A = \{G_1, G_2, \dots, G_n\}$. Каждой из дуг $G_j(e_i, e_k)$ соответствуют переходы D_i для пары точек p_k и p_i .

Каждая стратегия начинается с *изначального перехода* D_0 , осуществив который считаем, что искомое понятие может принадлежать данной категории. Изначальный переход ведет к *изначальной точке* e_k , соответствующей изначальному понятию e_0 категории Kat . Пример соответствия категории и дерева решений изображено на (Рис. 8).

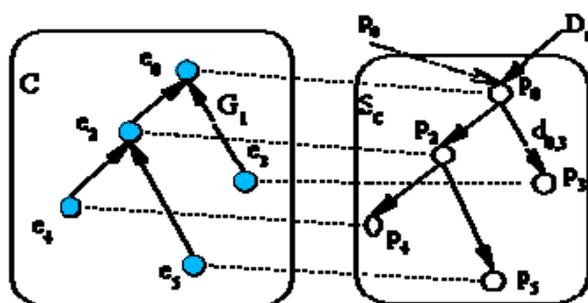


Рисунок 8 – Стратегия выбора понятия

В данном примере стратегия S_c построена для категории Kat . Каждому из классов понятий e_0, e_2, e_3, e_4, e_5 соответствуют точки p_0, p_2, p_3, p_4, p_5 , каждой из дуг G_j множества A соответствует определенная группа дуг D_j из множества D . Например, на рисунке отношению наследования $G_1(e_3, e_0)$ соответствует переход $d_{0,3}$. Из точки p_0 в точку p_2 и из точки p_2 в p_5 имеется несколько переходов – 2 и 3 соответственно.

Движение по графу стратегии S_c начинается с изначального перехода D_0 успешность которого ведет к изначальной точке p_0 . Далее проверяется возможность осуществления переходов D_2 или D_3 .

Способность осуществить такой переход определяется истинностью решающей правила эталонного понятия для проблемного понятия. Для примера, изображенного на (Рис.5), таким переходом является $d_{0,3}$. Например, если категория Kat содержит типы организационных структур, состоящую из различного рода организаций и их подразделений, переход $d_{0,3}$ требует наличия в проблемном понятии следующего признака: «Организация управляется коллективом спасателей».

Дальнейшее продвижение вниз по иерархии сужает критерии поиска эталонного понятия. Понятие считается найденным, если продвижение на нижние уровни с точки зрения предметной области не имеет смысла. Если проблемное понятие не может быть классифицировано ни по одному из имеющихся в категории классов, то необходимо добавить новый класс понятий, взяв в качестве родительского понятие, за которым поиск теряет смысл. Будем считать, что поиск теряет смысл за определенным эталонным понятием, если:

1. У эталонного понятия в иерархии отсутствуют дочерние понятия.
2. Все дочерние понятия эталонного понятия обладают одним или более признаком, которого нет у проблемного понятия.

Из множества полученных понятий, за пределами которого поиск теряет смысл, выбирается то, с который проблемное понятие имеет наибольшее количество общих наиболее влияющих признаков на появление ЧС.

Рассмотрим в общем виде алгоритм категорийной классификации понятий. Примем за e_p понятие, которое классифицируется. В качестве категории – Kat принимаем категорию, для которой происходит идентификация понятий. Экспертом будем считать некоторое действующее лицо, которое способно положительно или отрицательно ответить на вопрос о наличии определенного признака у классифицируемого понятия. В роли эксперта может выступать человек или ЭС.

Шаг 1. Эксперт принимает решение о возможности перехода D_0 .

Шаг 2. Если переход невозможен, то считается, что понятие e_p не может принадлежать категории Kat и процесс классификации в данной категории прекращается.

Шаг 3. Переходим к следующей точке p_i , в которую приводит нас разрешенный положительно переход D_i .

Шаг 4. Получаем множество возможных переходов $D_k, k \in \{k_1, k_2, \dots, k_m\}$, где m - количество дочерних классов для класса понятий e_i в категории Kat .

Шаг 5. Если число переходов равно нулю, то запоминаем класс, соответствующий точке текущей p_i в списке кандидатов Can родительских понятий для классифицируемого e_p . Возвращаемся в точку, из которой пришли в точку p_i . Если точка возврата p_0 , то заканчиваем поиск.

Шаг 6. Проверяем возможность осуществления каждого из переходов, полученных на шаге 4.

Шаг 7. Для каждого из возможных переходов продолжаем работу начиная с пункта 3 и возвращаясь на шаг 7 после завершения просмотра.

Шаг 8. По окончании проверки возвращаемся в ту точку, из которой пришли в точку p_i . Если точкой возврата является p_0 , то заканчиваем поиск.

Алгоритм идентификации понятий включает в себя три этапа:

- Поиск кандидатов.
- Выбор наиболее приближенного кандидата из категории.
- Выбор категории, к которой относится понятие.

Попытка классификации в обычном режиме проводится для всех категорий, имеющих в группе категорий или по выбору эксперта, на основании состояния предметной области, поиск может быть ограничен категориями по его выбору.

Наиболее подходящим считается класс, который имеет максимальное количество общих признаков с классифицируемым понятием. Если в категории обнаружено несколько классов кандидатов с одинаковым количеством общих признаков, то родительский класс выбирается либо по усмотрению эксперта либо любой (первый) из списка.

Затем сравниваются кандидаты из различных категорий и из них выбирается один с максимальным количеством общих признаков. В случае, если окажется несколько кандидатов с одинаковым максимальным количеством общих признаков, то решение о принадлежности той или иной категории принимает эксперт на основании собственного опыта.

Как было представлено ранее ситуация состоит из множества микроситуаций. В общем виде ситуацию предлагается представлять следующим образом:

- объявление объектов понятий;
- описание свойство объектов понятий;
- описание микроситуаций.

Объявление объектов понятий требуется, когда в ситуации представлены два или более однотипных понятия, например, одинаковое оборудование или сотрудники, занимающие одну и ту же должность в организации и(или) выполняющие одинаковые роли в данной ситуации.

Если требуется каким-либо образом охарактеризовать объекты понятий, то для них указываются значения атрибутов.

Каждую микроситуацию можно описать в виде простого предложения, состоящего из следующих частей – подлежащее, сказуемое и дополнение. Эти составляющие описывают центральное понятие, контекст и второстепенные понятия соответственно.

Данные принципы были положены в основу языка представления ситуаций, который позволяет описать проблемную ситуацию и затем провести поиск эталонной ситуации в базе знаний и выбрать связанное с ней проектное решение.

Состав описания ситуации выглядит следующим образом:

*ситуация : группа-микроситуаций | объявление-понятий группа-микроситуаций |
объявление-понятий свойства-понятий группа-микроситуаций*

Для сравнения ситуаций была определена мера сходства, выражающая расстояния между ситуациями. Сравнение ситуаций проводится путем расчета расстояний между микроситуациями, который включает в себя подсчет коэффициента подобия для входящих в данную микроситуацию понятий.

Общий принцип подсчета расстояний между понятиями, микроситуациями и ситуациями – это выделение общей части сравниваемых объектов и вычисление того, насколько каждый из них отличается от общей части.

Выводы

В результате проведенных исследований была сформирована методология микроситуационного представления ЧС, позволяющая эффективно применять повторное использование для поиска управляющих решений для ликвидации последствий ЧС. Была разработана технология, которая позволяет получить результат с учетом количественных и качественных параметров в моделях ЧС. Данное исследование дает возможность в условиях многокритериальности и неопределенности получить эффективное решение за счет использования прецедентных БД и БЗ. Следует подчеркнуть, что здесь была разработана технология наполнения прецедентную БЗ для ЭС и СППР за счет поиска информации в Интернет и социальных сетях. Развитие предлагаемого исследования будет направлено на разработку программных средств для ЭС и СППР в СЦ.

Используемые литературные источники

1. M.Ayaz Ahmad, Oleksandr Kuzomin, Vyacheslav Lyashenko, N. Ameer Ahamad. Microsituations as part of the formalization of avalanche climate to avalanche riskiness and avalanche safety classes in the emergency situations separation. ISSN 2320 5407 International Journal of Advanced Research (2015), Volume 3, Issue 4. P.684 -691.
 2. Oleksandr Kuzomin, Bohdan Tkachenko. Intellectual search engine of adequate information in internet for creating databases and knowledge bases. International Journal "Information Technologies Knowledge" Volume 9, Number 1, 2015. P.131- 138.
 3. Oleksandr Kuzomin, Mariia Tkachenko. Social search engine and intellectual database of people. International Journal "Information Technologies Knowledge" Volume 9, Number 1, 2015. P. 139 – 145.
-

Authors' Information



Oleksandr Kuzomin – Prof. Dr., Department of Informatics, Kharkiv National University of Radio Electronics; Kharkiv, Ukraine;

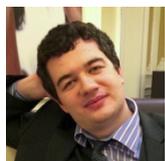
tel.: +38(057)7021515; e-mail: oleksandr.kuzomin@gmail.com

Major Fields of Scientific Research: General theoretical information research, Decision Making, Emergency Prevention, Data Mining, Business Informatics



Mihail Shtukin – PhD degree student, Kharkiv National University of Radio Electronics; Kharkiv; Ukraine; e-mail: shtumi@tut.by

Major Fields of Scientific Research: General theoretical information research, Decision Making, Emergency Prevention, Data Mining, Business Informatics



Bohdan Tkachenko – PhD degree student, Kharkiv National University of Radio Electronics; Kharkiv; Ukraine; e-mail: bohdan@tkachenko.io

Major Fields of Scientific Research: General theoretical information research, Knowledge Discovery and Engineering, Data Mining, Business Informatics



Mariia Tkachenko – PhD degree student, Kharkiv National University of Radio Electronics; Kharkiv; Ukraine; e-mail: mariia@tkachenko.io

Major Fields of Scientific Research: General theoretical information research, Knowledge Discovery and Engineering, Data Mining, Business Informatics

Information support of precedential approach in decision-making systems during emergencies

Oleksandr Ya. Kuzomin, Mihail Shtukin, Bohdan Tkachenko, Mariia Tkachenko

Abstract: *In this work we described database for emergencies prediction and decision making system, also we designed intellectual knowledge base for real-time data population and intellectual analysis*

Keywords: *emergencies, natural disaster, knowledge base, database, intellectual system, decision making*

ACM Classification Keywords: *H.3.3 Information Search and Retrieval*

Conference topic: *Information Technologies and Software Engineering - Information Systems*