

Экспертные системы

Экспертные системы - разновидность систем распознавания.

В экспертных системах процесс распознавания совмещен с процессом получения информации об объекте, причем последовательность получения значений признаков меняется в зависимости от полученных ранее данных о предыдущих признаках.

Экспертная система может отнести распознаваемый образ к известному ей классу до получения значений **всех** возможных признаков.

По обрабатываемым признакам экспертные системы можно разделить на системы с детерминированными и вероятностными признаками (возможно и их сочетание)

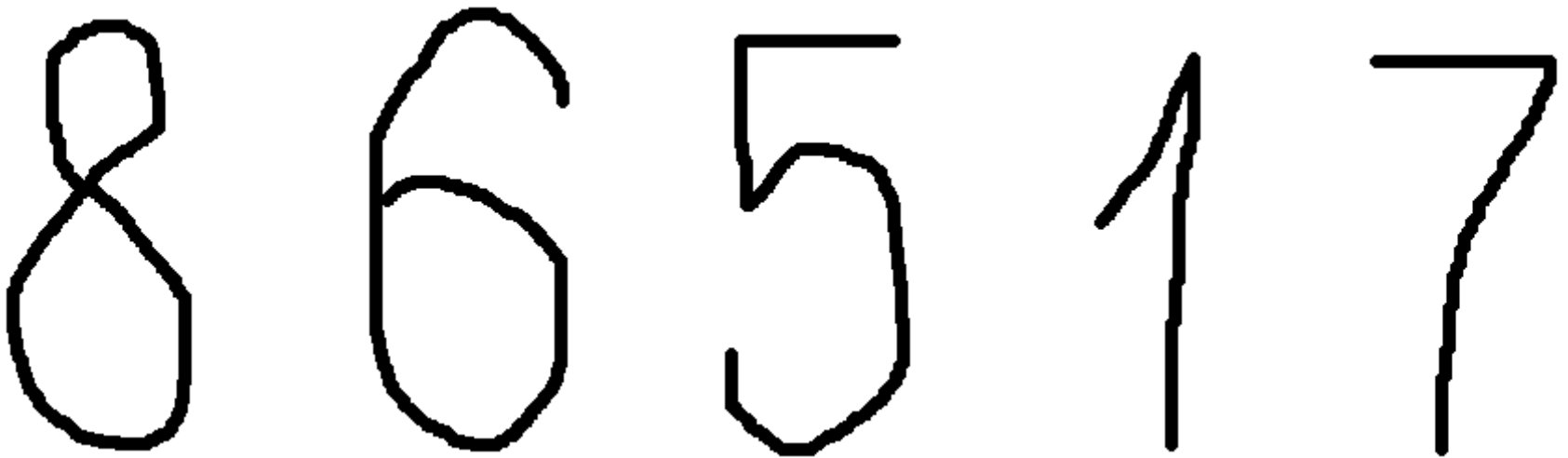
Система с детерминированными признаками

Разработчик строит дерево решений, с каждым узлом которого связана функция выбора, которая определяет – какой дочерний узел следует обрабатывать далее. Каждый листовой узел связан с конкретным классом, распознаваемым системой.

Пример дерева решений

Построение дерева решений для распознавания изображений рукописных цифр на основе следующих признаков:

- число замкнутых областей (озера);
- число незамкнутых областей (заливы);
- расположение залива (в верхней или нижней части цифры);
- направление открытой части залива (четыре возможные стороны).



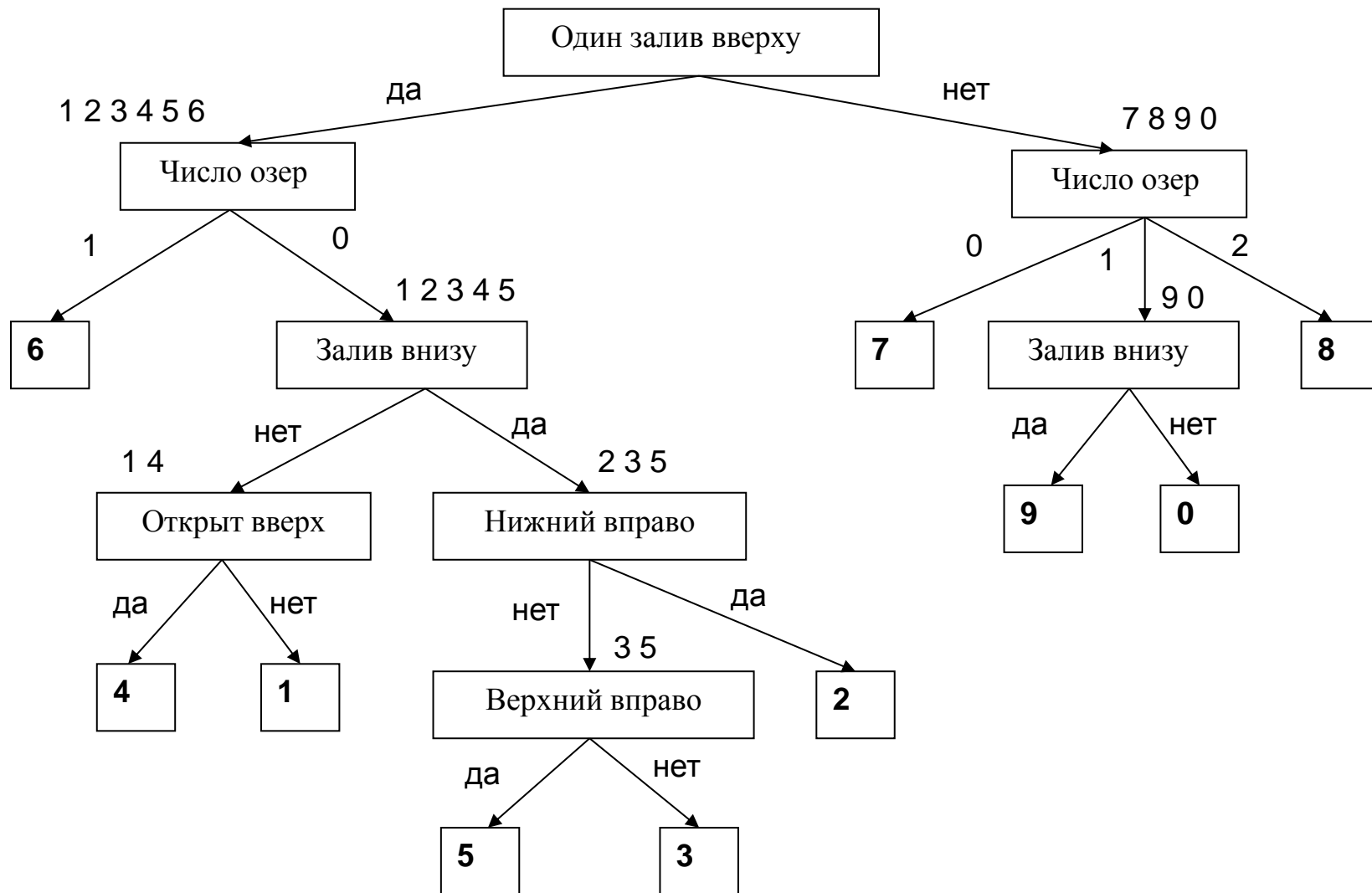
8 – два озера; 6 – одно озеро, один залив вверху, открытый вправо;
5 – два залива, один вверху, открыт вправо, один внизу открыт влево;
1 – один залив вверху открыт вниз; 7 – нет озер и заливов.

Признаки рукописных цифр

Цифра	Число озер	Число заливов	Расположение заливов	Направление
0	1	-	-	-
1	-	1	верх	вниз
2	-	2	верх, низ	влево, вправо
3	-	2	верх, низ	влево, влево,
4	-	1	верх	вверх
5	-	2	верх, низ	вправо, влево
6	1	1	верх	вправо
7	-	-	-	-
8	2	-	-	-
9	1	1	низ	влево

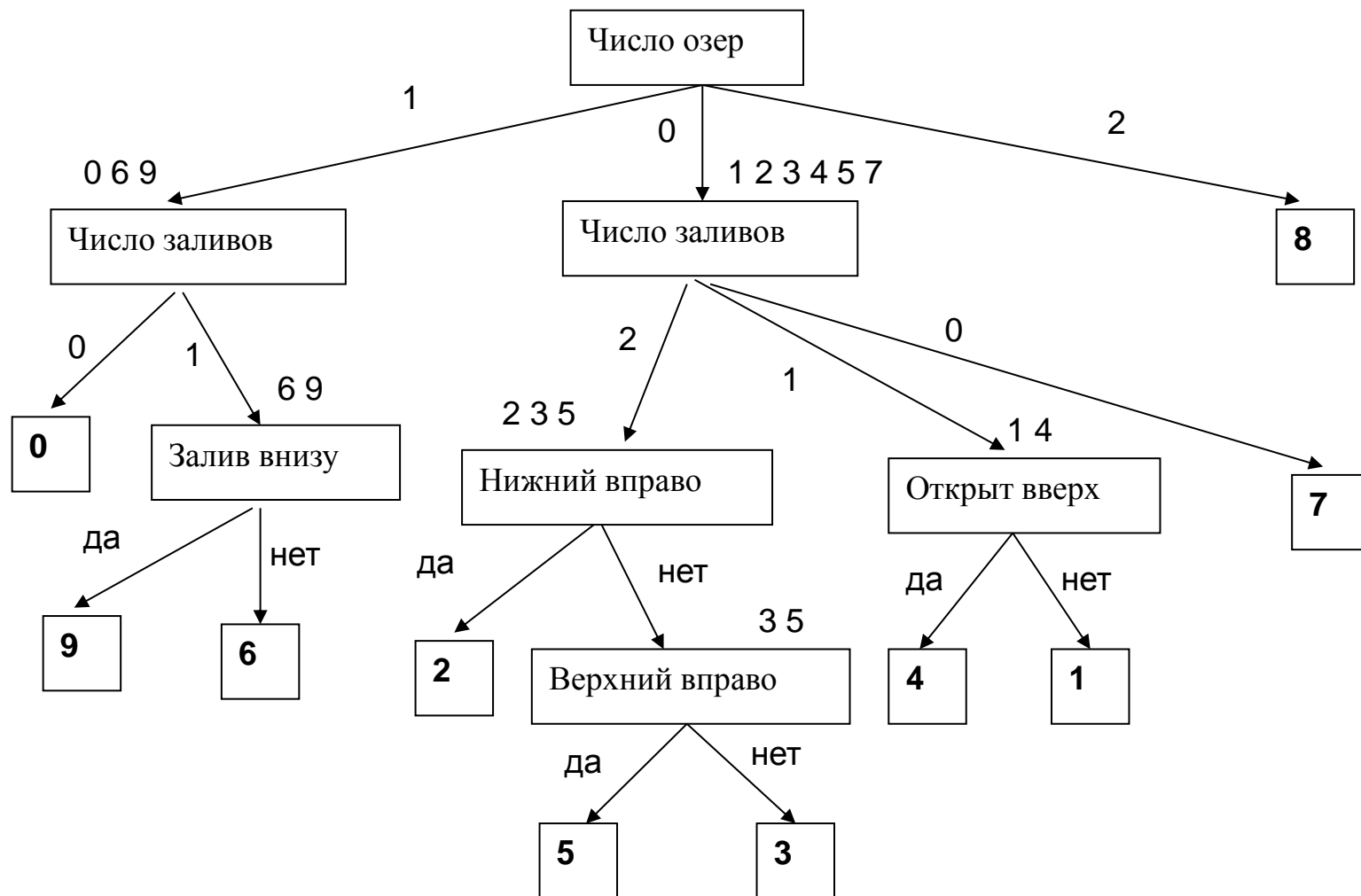
Замечание. Набор признаков не претендует на оптимальность, но распознавание по нему возможно, т.к. нет двух классов с одинаковыми наборами значений признаков.

Дерево решений



Цифра «6» классифицируется после второго вопроса. Максимально нужно 5 вопросов.

Другое дерево решений



Цифра «8» классифицируется после первого вопроса. Максимально нужно 4 вопроса.

Замечание. На одном уровне запрос признака на разных ветвях может отличаться.

Экспертная система с вероятностными признаками

Экспертная система состоит из двух достаточно независимых частей – **программы-оболочки** для хранения и обработки информации и **базы данных** с необходимыми для работы системы сведениями о конкретной области знаний. Рассмотренная далее система достаточно универсальна, т.к. способна диагностировать объект в некоторой области знаний после заполнения базы данных системы соответствующей информацией. Система последовательно задает вопросы о состоянии объекта, причем их количество и порядок могут меняться в зависимости от полученных ранее ответов, и после анализа имеющихся данных принимает решение, т.е. относит объект к одному из известных системе классов.

Этапы разработки экспертной системы:

разработка структуры базы данных; разработка алгоритма выбора последовательности запросов и принятия решения; разработка диалогового интерфейса; заполнение базы данных; тестирование.

Структура базы данных

Задача - формализация представления знаний о связях между признаками (далее симптомы) и классами образов (диагнозами), при условии, что **связи носят вероятностный характер**, т.е. одно и тоже значение симптома может с разной вероятностью соответствовать нескольким диагнозам.

1.Список симптомов $\{x_i\}$, $i = 1 \dots n$, где n – число симптомов

2.Диагнозы $\{w_j\}$, $j = 1 \dots m$, где m – число диагнозов

3.Связь диагнозов с симптомами в виде таблицы:

W	P_a	x_1		x_i		x_n	
w_j	$P_a(w_j)$	$p(x_1/w_j)$	$p(x_1/now_j)$	$p(x_i/w_j)$	$p(x_i/now_j)$	$p(x_n/w_j)$	$p(x_n/now_j)$

$p_a(w_j)$ - априорная вероятность диагноза w_j

$p(x/w_j)$ - условная вероятность наличия симптома x_i при диагнозе w_j

$p(x/now_j)$ - условная вероятность наличия симптома x_i при всех других диагнозах

Важно: в строке могут отсутствовать условные вероятности для некоторых симптомов.

Алгоритм работы системы

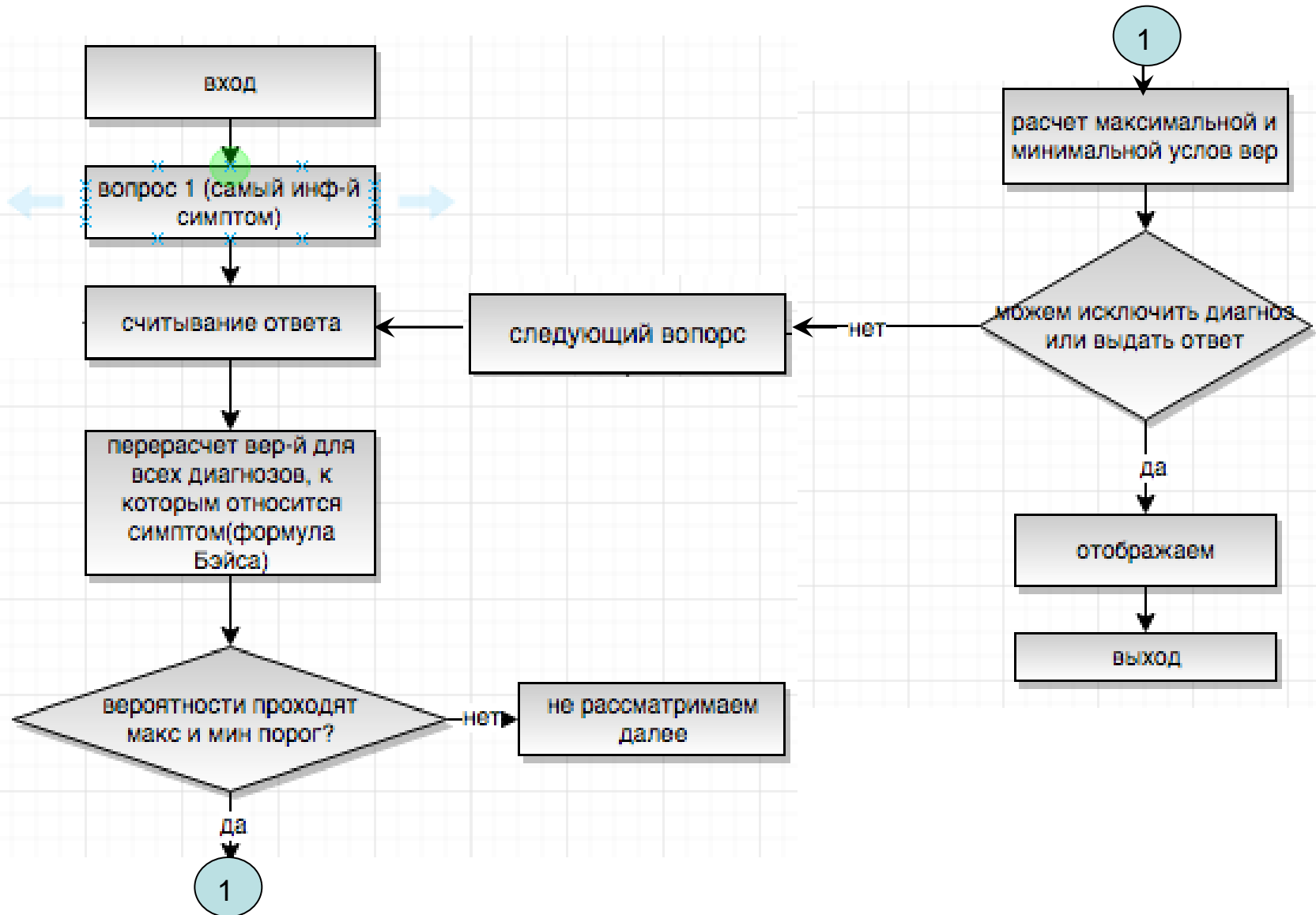
На каждом шаге:

1. Выбирается наиболее информативный симптом и задается соответствующий ему вопрос;
2. Полученный ответ используется для вычисления вероятностей диагнозов, и вопрос (симптом) исключается из дальнейшего рассмотрения;
3. По результатам вычисления некоторые диагнозы могут исключаться;
4. Если вероятность какого-то диагноза достаточна для принятия решения, то система заканчивает работу, если нет, то переход к п.1

Наиболее информативный симптом тот, информация о котором изменит вероятность максимального числа диагнозов (определяется по таблице связей диагнозов с симптомами)

Ответ (значение симптома) может быть бинарным (твердое «ДА» или твердое «НЕТ») или ответ может изменяться в некотором заданном разработчиком диапазоне, например, ± 5 , где -5 – твердое «НЕТ», +5 – твердое «ДА», 0 – отсутствие информации, остальные значения показывают степень уверенности (+1...+4) или неуверенности (-4...-1).

Алгоритм работы системы



Изменение вероятностей

Значение симптома используется для изменения вероятностей тех диагнозов, в описании которых он присутствует.

При твердом «**ДА**» вероятность диагноза w_j определятся по формуле Байеса:

$$p(w_j / x_i) = \frac{p(w_j) p(x_i / w_j)}{p(w_j) p(x_i / w_j) + (1 - p(w_j)) p(x_i / \text{now}_j)},$$

где $p(w_j)$ – вероятность диагноза w_j на текущем шаге.

При твердом «**НЕТ**» вероятность диагноза w_j определятся как:

$$p(w_j / x_i) = \frac{p(w_j)(1 - p(x_i / w_j))}{p(w_j)(1 - p(x_i / w_j)) + (1 - p(w_j))(1 - p(x_i / \text{now}_j))}.$$

До получения ответа на первый вопрос $p(w_j) = p_a(w_j)$, а после обработки ответа на заданный на данном шаге вопрос x_i будет $p(w_j) = p(w_j / x_i)$.

Изменение вероятностей

Учесть возможность получения ответа x_i в **заданном диапазоне** (± 5) можно линейной аппроксимацией (x_i берется со своим знаком).

Если:

$$x_i > 0 \quad p(w_j / x_i) = p(w_j) \left(1 + \left(\frac{p(x_i / w_j)}{p(x_i)} - 1 \right) \frac{x_i}{5} \right),$$

$$x_i < 0 \quad p(w_j / x_i) = p(w_j) \left(1 + \left(p(x_i / w_j) - \frac{p(x_i)(1 - p(x_i / w_j))}{1 - p(x_i)} \right) \frac{x_i}{5} \right),$$

$$x_i = 0 \quad p(w_j / x_i) = p(w_j),$$

где $p(x_i) = p(w_j)p(x_i/w_j) + (1 - p(w_j))p(x_i/\text{now}_j)$ – вероятность любого положительного ответа на вопрос x_i .

Оценка вероятностей

Если вероятность $p(w_j/x_i)$ близка к нулю, то диагноз w_j можно исключить из дальнейшего рассмотрения. Если вероятность $p(w_j/x_i)$ близка к единице, то диагноз w_j можно считать истинным и закончить работу системы. Пороги устанавливаются заранее и их точное значение подбирается экспериментально в процессе настройки системы.

Оценка возможных максимальных и минимальных вероятностей выполняется при условии, что симптомы статистически независимы.

Два симптома x_1 и x_2 статистически независимы, если для диагноза w_j условная вероятность $p(x_1 \& x_2/w_j) = p(x_1/w_j) p(x_2/w_j)$. Например, пусть из 100 автомобилей у 20-ти не горят фары, а у 10-ти не заводится двигатель. Эти события можно считать статистически независимыми, если у 2-х автомобилей одновременно не горят фары и не заводится двигатель ($20/100 * 10/100 = 2/100$)

У каждого диагноза просматривается список еще не рассмотренных симптомов и по каждому из них выполняется последовательность действий, приведенная далее (при первом цикле она выполняется для всех диагнозов).

Оценка вероятностей

Перед началом работы определяется какой ответ – положительный или отрицательный, увеличивает вероятность диагноза w_j .

Если $p(x/w_j) > p(x/\text{now}_j)$, то положительный ответ увеличивает $p(w_j/x_i)$ и при оценке возможных вероятностей диагноза расчетные условные вероятности $p'(x/w_j)$ и $p'(x/\text{now}_j)$ не меняются, т.е. $p'(x/w_j) = p(x/w_j)$,
 $p'(x/\text{now}_j) = p(x/\text{now}_j)$.

Если же, наоборот, $p(x/w_j) < p(x/\text{now}_j)$, то $p(w_j/x_i)$ увеличивается при отрицательном ответе и для оценки следует использовать обратные значения условных вероятностей, т.е. считать $p'(x/w_j) = 1 - p(x/w_j)$,
 $p'(x/\text{now}_j) = 1 - p(x/\text{now}_j)$

Если симптомы статистически независимы, то максимально и минимально возможные вероятности диагноза можно вычислять как произведение условных вероятностей по всем еще не полученным симптомам данного диагноза, при условии соответственно только положительных и/или отрицательных ответов – таких ответов, которые **максимально** увеличивают или уменьшают вероятность диагноза.

Оценка вероятностей

Вычисляются четыре условные вероятности (здесь K – число еще не рассмотренных симптомов диагноза w_j):

- вероятность диагноза w_j
при всех положительных ответах

$$p(w_j / x_{\Pi}) = \prod_{i=1}^K p'(x_i / w_j);$$
- вероятность не диагноза w_j
при всех положительных ответах

$$p(\text{now}_j / x_{\Pi}) = \prod_{i=1}^K p'(x_i / \text{now}_j);$$
- вероятность диагноза w_j
при всех отрицательных ответах

$$p(w_j / \text{nox}_{\Pi}) = \prod_{i=1}^K (1 - p'(x_i / w_j));$$
- вероятность не диагноза w_j
при всех отрицательных ответах

$$p(\text{now}_j / \text{nox}_{\Pi}) = \prod_{i=1}^K (1 - p'(x_i / \text{now}_j)).$$

Оценка вероятностей

Далее определяются **максимально возможная** вероятность диагноза w_j :

$$p_{\max}(w_j) = \frac{p(w_j)p(w_j / x_{\Pi})}{p(w_j)p(w_j / x_{\Pi}) + (1 - p(w_j))p(\text{now}_j / x_{\Pi})}$$

и **минимально возможная** вероятность диагноза w_j :

$$p_{\min}(w_j) = \frac{p(w_j)p(w_j / \text{nox}_{\Pi})}{p(w_j)p(w_j / \text{nox}_{\Pi}) + (1 - p(w_j))p(\text{now}_j / \text{nox}_{\Pi})}.$$

Если $p_{\max}(w_j) < p(w_j)$, то диагноз w_j можно исключить из рассмотрения.

Поиск наиболее вероятного диагноза.

Находим диагноз w_d , у которого $p_{\min}(w_d) \geq p_{\min}(w_j)$ для всех $j = 1 \dots k$, где k – число еще не исключенных диагнозов. Если $p_{\min}(w_d) \geq p_{\max}(w_j)$ для всех $j = 1 \dots k$ ($j \neq d$), то диагноз w_d считается наиболее вероятным и сообщается пользователю. Если нет, то цикл повторяется...

Пример работы системы

Исходные значения $p_a(w_j)$, $p(x/w_j)$, $p(x/\text{now}_j)$

	Диагноз 1	Диагноз 2	Диагноз 3
$p_a(w_j)$	0,3	0,2	0,1
Симптом 1	0,6/0,01	-	0,8/0,03
Симптом 2	0,7/0,02	0,8/0,01	0,02/0,1
Симптом 3	0,1/0,2	0,9/0,02	0,9/0,01
Симптом 4	-	0,01/0,07	0,7/0,04

$n=4$ ($i=1\dots 4$), $m = 3$ ($j=1\dots 3$). Пусть симптомы статистически независимые. Информация неполная, т.к. некоторые условные вероятности отсутствуют.

Найдем условные вероятности диагноза 1 до получения значения какого-либо симптома.

$$p(w_1/x_{\Pi}) = 0,6 \cdot 0,7 \cdot 0,9 = 0,378; \quad p(\text{now}_1/x_{\Pi}) = 0,01 \cdot 0,02 \cdot 0,8 = 0,00016;$$

$$p(w_1/\text{no}x_{\Pi}) = 0,4 \cdot 0,3 \cdot 0,1 = 0,012; \quad p(\text{now}_1/\text{no}x_{\Pi}) = 0,99 \cdot 0,98 \cdot 0,2 = 0,19404.$$

Замечание.

У симптома 3 используются обратные условные вероятности, т.к. $p(x_3/w_1) < p(x_3/\text{now}_1)$

Далее найдем максимально и минимально возможные вероятности диагноза 1

$$p_{\max}(w_1) = (0,3 \cdot 0,378) / (0,3 \cdot 0,378 + 0,7 \cdot 0,00016) = 0,999;$$

$$p_{\min}(w_1) = (0,3 \cdot 0,012) / (0,3 \cdot 0,012 + 0,7 \cdot 0,19404) = 0,02582.$$

Пример работы системы

Начальные p_{\max} и p_{\min}

	Диагноз 1	Диагноз 2	Диагноз 3
$p(w_j/x_{\Pi})$	0,378	0,7128	0,49392
$p(now_j/x_{\Pi})$	0,00016	0,00019	0,00119
$p(w_j/nox_{\Pi})$	0,012	0,0002	0,00012
$p(now_j/nox_{\Pi})$	0,19404	0,06791	0,09219
$p_{\max}(w_j)$	0,99901	0,99893	0,97878
$p_{\min}(w_j)$	0,02582	0,00074	0,00014

Информация о симптомах 2 и 3 присутствует во всех диагнозах. Выбираем первым вопрос о симптоме 2. Пусть ответ – твердое «ДА» ($x_2 = 5$).

Текущие значения вероятностей диагнозов при учете информации о симптоме 2:

$$p(w_1/x_2) = (0,3 \cdot 0,7) / (0,3 \cdot 0,7 + 0,7 \cdot 0,02) = 0,9375;$$

$$p(w_2/x_2) = (0,2 \cdot 0,8) / (0,2 \cdot 0,8 + 0,8 \cdot 0,01) = 0,95238;$$

$$p(w_3/x_2) = (0,1 \cdot 0,02) / (0,1 \cdot 0,02 + 0,9 \cdot 0,1) = 0,02174.$$

Примем порог достоверности 0,99. Наиболее вероятен диагноз 2, но $p(w_2/x_2) < 0,99$

Пример работы системы

$p(w_j)$, p_{\max} , p_{\min} после ответа на вопрос 2

	Диагноз 1	Диагноз 2	Диагноз 3
$p(w_j)$	0,9375	0,95238	0,02174
$p(w_j/x_{\Pi})$	0,54	0,891	0,504
$p(now_j/x_{\Pi})$	0,008	0,0186	0,00001
$p(w_j/nox_{\Pi})$	0,04	0,001	0,006
$p(now_j/nox_{\Pi})$	0,198	0,0686	0,92189
$p_{\max}(w_j)$	0,99901	0,99896	0,99893
$p_{\min}(w_j)$	0,75188	0,22573	0,00015

Ни один диагноз нельзя исключить. Максимальное значение p_{\min} имеет первый диагноз, но $p_{\max}(w_2) > p_{\min}(w_1)$, т.е. первый диагноз пока нельзя считать наиболее вероятным.

Из трех оставшихся симптомов наиболее информативный – симптом 3 (связан со всеми диагнозами). Пусть на него получен ответ – твердое «НЕТ» ($x_3 = -5$).

Пример работы системы

$p(w_j)$, p_{\max} , p_{\min} после ответа на вопрос 3

	Диагноз 1	Диагноз 2	Диагноз 3
$p(w_j)$	0,94406	0,67114	0,00022
$p(w_j/x_{\Pi})$	0,6	0,99	0,56
$p(now_j/x_{\Pi})$	0,01	0,93	0,012
$p(w_j/nox_{\Pi})$	0,4	0,01	0,06
$p(now_j/nox_{\Pi})$	0,99	0,07	0,9312
$p_{\max}(w_j)$	0,99901	0,68479	0,51160
$p_{\min}(w_j)$	0,87209	0,22573	0,00015

Наиболее вероятен диагноз 1, но $p(w_1) < 0,99$.

Максимальное значение p_{\min} имеет диагноз 1. $p_{\max}(w_2) < p_{\min}(w_1)$ и $p_{\max}(w_3) < p_{\min}(w_1)$, следовательно, **диагноз 1 можно считать наиболее вероятным**. Работа закончена.

Замечание. Значение двух симптомов (1 и 4) не потребовалось.

Если на каждый вопрос отвечать (+5), то система задаст **все** вопросы и выберет в качестве наиболее вероятного диагноз 3, т.к. после окончания работы $p(w_3) = 0,99893$. $p(w_2) = 0,99228$ – почти совпадает с $p(w_3)$, а $p(w_1) = 0,97826$, что меньше порога 0,99.