Первоисточник: [http://fuzzy.kstu.ru/fulltext.htm]

Вероятность следствия, Логическая версия теоремы Байеса, и Операции нечеткой логики

Интуитивно, когда мы говорим, что следствие*, "A* если *B" (A* « - *B)* верно, мы подразумевает, что всякий раз, когда *B* истина, то мы можем заключить, что *A* тоже истина. Другими словами, следствие - то, что позволяет нам сделать логический вывод.

Во многих практических ситуациях, мы имеем немного веры{секретности} *B,* но мы не на 100 % уверены, что *B* является верным. Точно так же мы не можем быть уверенных 100 %, что значение *A* « - *B* является верным. В таких ситуациях, мы можем оценить вероятность *P {B)*, что *B* является верным, и вероятность *P (" я " B)*, что значение *A* ч-*B* верно. Как мы можем исполнить логический вывод в таких ситуациях? Интуитивно, мы ожидаем быть способными заключить, что в этом случае, должен также быть верен с некоторой вероятностью; эта вероятность должна иметь тенденцию к 1 как вероятности *P {B)*, и *P (" я " B)* имею тенденцию к 1.

Как мы можем расширить логическое значение на вероятностный случай? В зависимости от того, как мы интерпретируем вероятность значения, мы становимся два отличными

Есть два известных ответа на этот вопрос, и эти ответы отличны, потому что они используют различную формализацию вероятности значения. Первый ответ от подхода Байеса, в котором *P (4-B)* интерпретируется как условная вероятность *P (A\B)*. Второй ответ прибывает от логического рассуждения, где вероятность *P (<-B)* интерпретируется как вероятность передачи "материальное значение", то есть, вероятность *P (A V notB)*, что или является верным, или *B* ложен.

От просто логической точки зрения, второй ответ может звучать более разумным, но есть примеры, где первый ответ находится в лучшем соответствии со здравым смыслом. Действительно, предположите, что мы анализируем животных в национальном парке, и мы ищем вероятность значения *A <- B,* где - "животное, белое ", и *B* - "животное - тигр". В простом английском языке, вероятность утверждения*, "*если *B"* естественно интерпретируется как вероятность, что тигр является белым. Если из 10 000 животных, 100 - тигры, и 10 из этих тигров белые, то, в условиях здравого смысла, вероятность, что тигр белый - 10/100=0.1. Это - точно вероятность, обеспеченная подходом Байеса. Однако, логический подход производит различный результат: вероятность V-iB, то есть., вероятность, что животное является или белым или не тигр, равна 9,910/10,000=0.991 - потому что утверждение*,* V-> *B* верно не только для 10 белых тигров, но также и для 9 900 животных, которые - не тигры.

Эти примеры показывают, что нет ни одной "правильной" вероятностной интерпретации значения, но в зависимости от ситуации, различные интерпретации могут быть разумны. Поэтому желательно обеспечить сравнительный анализ различных интерпретаций.

Показано, что вышеупомянутые две интерпретации могут быть представлены как специфические случаи более общего подхода, в котором различие соответствует различию между различными t-norm-like действиями (для детальной информации относительно t-норм).

Теперь представим, что мы интерпретируем значение как материальное значение *A* V not(B) . В этом случае, вероятность значения интерпретируется как вероятность *P(A* V not(B)). Если мы знаем эту вероятность и если мы знаем *P (B),* как мы можем определить *P (A&B)?* Базируемый на аддитивности вероятности и факта, что и-**A* являются несовместимыми, мы заключаем что *P (B) = P (AkB) + P (-.AkB)*. Поэтому,

P (AkB) = P (B)-P (^AkB). (10)

Утверждение{Заявление}-**Ak B* - отрицание V-iB, следовательно

P (^AkB) = l-P (iVnB). (11)

Заменяя (11) в (10), мы заключаем это

P (AkB) = P (B) + P (AV ^B) - 1. (12)

Эта формула подобна формуле (2): обе формулы могут быть описаны как

(13)

P (КБ) = P (" я " B) © *P (B)*

для некоторого двойного действия ® *b.* В формуле (2) - который соответствует Bayesian случаю - мы использовали функцию ® *b = a-b*

В формуле (12) - который соответствует логическим случаям{делам} значения - мы использовали действие *a@b = + b - 1.* Так как значение действия © должно преобразовать вероятности в новую вероятность, и вероятности только берут ценности от интервала [0,1], разумно требовать, чтобы действие *aQb* всегда взяло ценности от интервала [0,1]. Действие *a@b* = + *b* - 1 не всегда удовлетворяет это требование, потому что, когда и *b* - оба, скажем, меньше чем 0.5, мы имеем + *b* - 1 <0. Это не затрагивает наше заявление{применение}, потому что мы всегда имеем *P {B)* + *P (*V-iB)> 1. Однако, чтобы делать действие © *b* всюду определенный как функция от вероятностей до вероятностей, разумно установить ее ценность на 0 когда + *b* - 1 <0, то есть, рассматривать{считать} новое действие *a@b* = Макс (0, + b - l).

Оба действия (14) и (15) - примеры *t-норм,* действия, описывающие "и" в нечеткой логике. Неофициально, появление t-нормы имеет смысл, потому что *к B* верен, если В верен*, и* значение-> *B* верно, так что разумно заключить, что наша степень{градус} веры *P (к B)* в *AkB* равна результату "и" - действие (t-норма) © *b* прикладной к степеням{градусам} веры *P (B)*, что *B* является верным и степень{градус} веры *P (« - B)*, что значение верно. Это оправдание неофициально. В следующем тексте, мы будем делать более формальное оправдание.

Тем временем, так как формула (12) подобна формуле (2), мы будем использовать эту аналогию, чтобы вывести логический аналог вывода формулы Заливов. С тех пор *AkB* означает ту же самую вещь как *B к A,* мы поэтому заключаем что *P (AkB) = P (BkA),* то есть, из-за (13), что

P (<-B) ® P (B) = P (B <-A) ® P (A). (16)

В частности для исчерпывающего списка *n* несовместимых гипотез *Привет...,* $_{Hn}$*,* и для свидетельства{очевидности} *E,* мы заключаем это

P {Щ <-E) ® P {E) =P {E <-Щ) ® P {Щ). (17) Поэтому,

P (Щ i-E) = (P (E " я " Щ) © P {Привет)) 0 *P (E),* (18)

где a0b - обратное действие к ©, то есть, действие для который (о 0 *b)* © *b = a.*

Это стоит упоминать, что для общей t-нормы ©, соответствующее обратное действие 0 обычно называют нечетким значением [2, 4].

Для умножения (14), обратное действие - разделение 0 *b* = *a/b* (как используется в формуле (6)). Чтобы удостовериться, что ценности этого действия остаются в пределах интервала [0,1], мы должны заменить это

о0 *b = мм (a/b,* 1).

Для нашего специфического действия (15), обратное действие - 0 *b* = минута (1+ - *b,* 1);

здесь, подобно к случаю разделения, мы добавили минуту (l...), чтобы удостовериться, что ценность этого действия остается в пределах интервала [0,1].

Из-за формулы (7) и факта это

*P (EкЩ) = P {E *-Щ) © P (Щ),* мы получаем выражение для *P {E):*

P (E) = P (E " я " Привет) ® P (Привет) +. + P (E " я " $_{Hn}$*) ® P* $_{(Hn)}$. Так, мы заканчиваемся что

P (Щ <-E) = (P (E <-Щ) © P (Щ)) 0 (19) *(P (E *-Привет) © P (Привет) + ... + p (*$_{E\wedge Hn}$*) ® P (*$_{Hn}$*)).* Это - прямой логический аналог теоремы Заливов.

Общий подход к описанию вероятности значенияи его отношение к нечеткой логике

Позвольте нам описывать общее определение вероятности *P {<-B).* Эта вероятность должна только зависеть от событий и *B.* Таким образом, наше первое требование - то, что, как только мы знаем вероятности всех возможных Булевых комбинаций и *B,* мы должны быть способны определить желательную вероятность *P (<-B).*

Известно, что, чтобы определять вероятности всех возможных Булевых комбинаций и *B,* достаточно знать вероятности *P (к B)*, *P (к*-> B), *P {-*AkB),* и *P {-^Ak-^ B)* четырех *атомных* утверждений{заявлений} *AkB, Ak-> B,-> AkB,* и-> *Ak-> B.* Для простоты, в следующем тексте, мы обозначим соответствующие вероятности *Рц,* Рю, Пои, и Роо-передача, четыре атомных утверждения составляют всесторонний список несовместимых событий, так что их сумма должна быть равна 1:

P **(*КБ)* + *P (к-> В)* + Р (-.Л *к В)* + *Р (-, к-> В)* = 1.**

Таким образом, мы можем определить общее действие значения как функция этих четырех вероятностей:

Определение 1. *Распределением вероятности Р, мы подразумеваем учетверенный из неотрицательных ценностей* Рц, Рю, Пои, и Роо*, для который Рц* + Рю + Пои + Роо = " я " *набор всех распределений вероятности обозначен V.*

Определение 2. *Вероятностным логическим действием, мы подразумеваем функцию F: V->* [0,1]. *Для каждых двух событий и В, результат применения{обращения} вероятностного логического действия F тогда определен как*

$$F(P) = F(P(A\&B), P(A\& notB), P(notA \& notB)) \qquad 11$$

Для определения Байеса (1), мы имеем $P(A\&B) = P_{11}$, и $P(B)$ = Р $(A \& B)$ + $P(notA \& B) = P_{11} + P_{01}$, следовательно f($P_{11}$, P_{10}, P_{01}, P_{11})= P_{11}/(P_{11}+ P_{11})

Для логического определения *Р (*ч-*В)* = *Р (*V-> *В),* мы имеем (из-за (11)) *Р (AV-> В) = l-P (^AkB)* = l-POb следовательно

F (Pu, Пьо, fl) я, Pii) = с 1 пои. (21)

Определение 1 - общее определение вероятностного логического действия, это не различает между значением, соединением, дизъюнкцией, и т.д. Что делает действие действие значения? Одна вещь, которая является верной для значения *4-В* а не для других действий - то, что значение зависит только от того, что случается, когда *В* верен

и нельзя затронуть тем, что случается, когда *В* ложен. Другими словами, если для двух распределений, мы имеем те же самые ценности *Р (AkB)* и *Р {-*AkB),* затем для этих двух распределений, мы должны получить точно ту же самую ценность *Р (Ай-В).*

Другое условие{состояние}, описывающее значение состоит в том, что, если, когда *В* всегда подразумевает *А,* то есть, когда-**AkB* невозможен (то есть, когда *Р {-> AkB)* =, ПОИ = 0), затем <-*В* должно быть верно с вероятностью 1.

Позвольте нам описывать эти условия формально:

Определение 3. Мы *говорим, что два распределения вероятности Р и Р' являются эквивалентными, когда В верен если Рц* = Pn *и* ■foi = -

политический •

Определение 4. Мы *говорим, что вероятностное логическое действие F - действие значения, если следующие два условия держатся:*

- *F (P) = F (P') для всех пар П и P', которые являются equiv*
- *alent, когда В верен;*
- *если* ПОИ = 0, *то F (P)* = 1.

Оба действия (20) и (21) - действия значения в этом смысле. Вообще, следующие простые суждения обеспечивают полные описания таких действий значения:

Утверждение 1. *Вероятностное логическое действие F - действие значения, если и только если F зависит только от двух переменных Рц andPOi, то есть, ifF (P)* = / (Pu, Пои) *для некоторой функции f двух переменных для который* / (Рц, 0) = 1 *за все ценности* Рц.

С тех пор

Пои = *РЫ к В) = Р {В) - Р {к В)* = Р*i - Рц,

где мы обозначили Р*i = *Р {В),* мы можем повторно сформулированное Суждение 1 следующим образом:

Утверждение 1'. *Вероятностное логическое действие F - действие значения, если и только если F зависит только от двух переменных Рц andP*i, то есть, если F {P) = г (Pu, P*i) для некоторого г функции двух переменных для который д {Рц, Рц)* = 1 *за все ценности* Рц.

Таким образом, чтобы описать все возможные действия значения, мы должны описать соответствующие функции двух переменных.

Так как мы рассматриваем вероятностную неуверенность, разумно рассмотреть не только индивидуальные события *А, ',* и т.д., но также и "соединения" (вероятностные комбинации) таких событий. Общая идея позади таких комбинаций - то, что мы берем лотерею с некоторой вероятностью *р* и затем выбираем, если лотерея преуспевает и′ иначе. Согласно теории вероятности, вероятность заканчивающегося случая равна *Р (А) =р-Р (А).*

И формула Байеса и логическое значение охвачены формулой (29):

- формула Байеса соответствует = 0, и
- логическая формула значения соответствует = 1.

Таким образом, наше заключение - то, что естественные требования определяют 1-параметрическое семейство формул для вероятности значения, формулы, которые являются промежуточными между двумя чрезвычайными случаями: Байеса и логичный.

Вообще, для *h (b)* = + (1 - *a)* ■ *b,* соответствующее действие скопления (28) имеет форму

t (a, b) = © *b* = ■ *(+ b* - 1) + (1 - *a)* ■ ■ *b.* (30)

Можно легко проверить, что это действие является всегда ассоциативным.

Для этого действия ©, обратное действие *c* = <*Z*> *b* может быть определено от уравнения = b © с, то есть, = ■ *(b* + с - 1) + (1 - *a)* ■ *b* ■ *c* Это - линейное уравнение в терминах с, от которого мы заключаем это.

Можно легко проверить, что для = 0, мы получаем инверсию Заливов *a/b,* и для = 1, мы получаем обратное действие 1 + - *b* передающий логическому значению.

Как мы можем интерпретировать эти новые действия скопления? Можно показать, что, если мы "повторно нормализуем" вероятности, используя преобразование

(31)

P-) • *s (P)* = (1 - а) ■ Р + а,

тогда в этом новом масштабе, действие скопления (30) становится простым изделием: *s (© b) = s (a)* ■ *s (b).*

Перевычисление (31) делает совершенный смыс. Действительно, один из естественных методов приписывать субъективную вероятность *P {A)* к утверждению состоит в том, чтобы взять несколько *(N)* экспертов, и спросить каждого из них, или он или она верят, что верен. Если *N (A)* их отвечают "на да", мы берем *d {A) = N (A)/N* как желательная ценность уверенности. Если все эксперты верят в *A,* то эта ценность - 1 (=100 %), если половина из них верит в *A,* то *t (A)* = 0.5 (50 %), и т.д.

Инженеры знания хотят, чтобы система включила знание полного научного сообщества, так что они спрашивают

так много экспертов насколько возможно. Но выяснение слишком многих экспертов ведет к следующему отрицательному явлению: когда мнение относительно наиболее уважаемых профессоров, победителей Нобелевского приза, и т.д., известно, некоторые менее уверенные в себе эксперты не будут достаточно храбры, чтобы выразить их собственные мнения, так что они будут скорее следовать за большинством. Как их присутствие влияет на заканчивающуюся субъективную вероятность?

Пусть N начальное число экспертов, *N (A)* число таковых из них, кто верит в *A,* и *M*- число сомневающихся экспертов. Первоначально, *d {A) = N (A)/N.* После того, как мы добавляем *M.* эксперты конформиста, число экспертов, которые верят в, становится *N (A) + M.* из общего количества *M. + N.* Так что новая ценность субъективной вероятности

$P'(A)=(N(A)+M)/(N+M)= (1-\alpha)*P(A)+ \alpha$

где $\alpha = M/N$.

Таким образом, каждое новое действие может просто интерпретироваться как действие Байеса, но в различном масштабе вероятности, естественном для опытных систем.