

АБДУКТИВНЫЕ РАССУЖДЕНИЯ В ЗАДАЧАХ ОБЪЯСНЕНИЯ НАБЛЮДАЕМОГО

© 2020 г. Академик РАН С. Н. Васильев^{1,*}

Поступило 01.06.2020 г.

После доработки 01.06.2020 г.

Принято к публикации 04.06.2020 г.

В проблематику искусственного интеллекта, управления и принятия решений при неполной или недостоверной информации входит широкий класс задач абдуктивного объяснения, в том числе задачи в терминах “причина—эффект”. Работа посвящена логическому формированию гипотез, объясняющих наблюдаемое. Предлагаются средства представления знаний и вывода гипотез. Вводится язык, обладающий свойством так называемой подстановочности. Свойства языка и вводимых в нем исчислений обеспечивают абдукцию путем сочетания дедукции с гипотезированием. В отличие от известных логических методов абдукции, предложенные средства обеспечивают получение гипотез (минорант), необходимых и достаточных для формального объяснения наблюдаемого. На основе минорант в сочетании с базовой теорией предметной области формируются достоверные причины наблюдаемого или находятся релевантные обстоятельства, выводящие на эти причины. При этом в ситуациях наличия также эмпирических данных эти причины и обстоятельства могут формироваться и в правдоподобных версиях. Рассматриваются примеры из техники и медицины.

Ключевые слова: абдуктивные рассуждения, правдоподобный вывод, объяснение наблюдаемого, диагностика, искусственный интеллект, интеллектуальное управление

DOI: 10.31857/S2686954320040190

ВВЕДЕНИЕ

В проблематику искусственного интеллекта (ИИ), управления и принятия решений при неполной или недостоверной информации входит широкий класс задач так называемого объяснения на основе моделей и методов абдуктивных рассуждений (Explanation by abductive reasoning), в том числе в терминах “причина—эффект”.

Понятие абдукции введено логиком и философом Ч.С. Пирсом (C.S. Peirce, 1839–1914) в форме логических рассуждений, сравнимых, но отличных от дедукции и индукции. Сегодня абдукция чаще всего определяется следующим образом: на основе заданной теории T предметной области и утверждения G о наблюдаемом эффекте, подлежащем объяснению, найти такую, не противоречащую теории T , гипотезу—объяснение Δ , что из T и Δ выводимо G (модификации см., например, в [1–3]).

В работах по абдуктивному гипотезированию на теорию T предметной области часто накладываются ограничения, при которых дедукция максимально упрощается, например, содержание теории должно

иметь форму “if-and-only-if”, т.е. форму семейства эквивалентностей $c_i \leftrightarrow E_{i1} \vee \dots \vee E_{in}$, где в пропозициональном случае c_i — переменные-причины, а E_{ij} , $j = 1, \dots, n$, — элементарные конъюнкции (ЭК), составленные из переменных-эффектов и их отрицаний, возможно, при дополнительных требованиях к множеству задействованных переменных, частично упорядоченному импликацией [1, 2].

В данной работе не предполагается какой-либо предварительной трансформации теории T , а в качестве языка представления знаний в разделе 1 вводится S -язык L_S как подмножество стандартного пропозиционального языка, но без ослабления его выразительной силы. Он шире пропозиционального варианта языка позитивно-образованных формул [4], допуская использование отрицательных литералов. Благодаря подстановочности формул языка L_S (см. раздел 1), они обладают свойствами, не все из которых справедливы для произвольных пропозициональных формул. Свойства языка L_S и вводимых в нем исчислений $\mathcal{C}_\alpha$, $\mathcal{C}_\beta$ обеспечивают удобное сочетание дедукции с гипотезированием.

Вывод в $\mathcal{C}_\beta$ наблюдаемого эффекта G выполняется вместе с синтезом формулы P , называемой гипотезой-минорантой и объясняющей эффект G лишь формально. С точностью до эквива-

¹ Институт проблем управления им. В.А. Трапезникова
Российской академии наук, Москва, Россия

*E-mail: vassilyev_sn@mail.ru

лентности гипотеза P логически минимальна (любая гипотеза-причина Δ , объясняющая G , должна быть не слабее $P: \Delta \rightarrow P$). Гипотеза-миноранта выступает контрольным условием и опорным средством формирования достоверных гипотез Δ , а в случае наличия сверхтеории T также эмпирических знаний, то и правдоподобных гипотез Δ . Эмпирические знания — это знания об априорном множестве потенциально возможных причин-кандидатов и/или о правдоподобных причинно-следственных связях. Из гипотезы P переходом к достаточным условиям, например, путем упрощения или использования эмпирических знаний, получаются либо причины Δ эффекта G (достоверные или правдоподобные), либо поначалу лишь релевантные обстоятельства (факты, события, “улики”), из которых в рамках теории T средствами исчислений C_α и C_β оказываются выводимыми и сами причины.

Заметим, что в проблематике интеллектуализации автономных (беспилотных) агентов с реактивными правилами “условие—действие” понятие “эффект” может иметь также смысл целевого состояния агента [5]. Эта тема абдуктивной информационной поддержки планирования действий агентов с распространением полученных в данной работе результатов на первопорядковые логики — актуальная задача на будущее, в части представления знаний, решаемая, как видится, с сохранением в языке свойства подстановочности позитивно-образованных формул [4, 6] и требований позитивности.

В разделе 1 вводится язык представления знаний, в разделе 2 — правила преобразования его формул. В разделе 3 вводятся исчисления C_α , C_β и рассматриваются их свойства, а в разделе 4 применение этих исчислений иллюстрируется примерами из техники и медицины.

1. ПРЕДСТАВЛЕНИЕ ЗНАНИЙ

В пропозициональном языке L стандартной (классической) семантики выделим подмножество L_S . Ввиду свойства подстановочности его формул, определяемой ниже, будем называть его S -языком (от английского Substitutability). Далее конъюнктами называем ЭК литералов (пропозициональных переменных и/или их отрицаний) и пропозициональные константы *false*, *true* (ЭК — это конъюнкции литералов; т.е. если в ЭК входит пара контрарных литералов (переменная и ее отрицание), то по умолчанию такие ЭК заменяются на константу *false*. Логические связи $\vee$, $\wedge$, $\neg$, $\rightarrow$ в S -языке понимаются стандартно.

Определение 1 (формулы языка L_S и их типы).

1) ЭК A или константа *false*, — S -формулы; им приписывается тип $\wedge$; это простейшие S -формулы;

2) если $F_i, i = 1, \dots, m$, — S -формулы типа $\wedge$, а G — ЭК или *true*, то $G \rightarrow \left(\bigvee_{i=1}^m F_i \right)$ — S -формулы типа $\rightarrow$;

3) если $F_i, i = 1, \dots, m$, — S -формулы типа $\rightarrow$, а G — ЭК или *true*, то $G \wedge \left(\bigwedge_{i=1}^m F_i \right)$ — S -формулы типа $\wedge$;

4) других формул в языке L_S нет.

Семантика S -формул очевидна. Их возможные упрощения в случае появления $F_i = \textit{false}$ (кроме случаев, когда $m = 1$ в формуле типа $\rightarrow$) также очевидны и такие упрощения будут подразумеваться. Формулы языка L вне L_S понимаем стандартно. Например, если $F_1, F_2 \in L_S$, то выражение $F_1 \rightarrow F_2$, вообще говоря, не принадлежащее L_S , понимается как в L .

Подформулы S -формулы F , образуемые в соответствии с индуктивным определением S -формул, называются главными подформулами (ГП). Выражение $F_1 \subseteq F$ означает, что F_1 — ГП S -формулы F . Каждая вершина древовидной структуры S -формулы имеет тип, соответствующий типу ГП, корнем которой эта вершина является. Типы $\wedge, \rightarrow$ вершин вдоль ветви структуры S -формулы F чередуются.

Для S -формул F, F_1, F_2 справедливо свойство подстановочности: если $F_1 \subseteq F$ и $F_2 \rightarrow F_1$, то $F(F_2/F_1) \rightarrow F$, где $F(F_2/F_1)$ — результат подстановки в F вместо F_1 не менее сильной формулы F_2 . Язык L_S полон относительно выразительной силы языка L .

2. ПРЕОБРАЗОВАНИЯ S -ФОРМУЛ

Каноническим представлением S -формулы F называем ее запись в виде формулы с конъюнктом *true* в корневой вершине ее структуры:

$$F = \textit{true} \rightarrow \bigvee_{i=1}^m \Omega_i, \quad \Omega_i = A_i \wedge \left(\bigwedge_{j=1}^{n_i} (B_{ij} \rightarrow \Phi_{ij}) \right), \quad (1)$$

где $m \geq 1, n_i \geq 0$ (если $n_i = 0$, то $\Omega_i = A_i$). Конъюнкт A_i называется базой, а конъюнкты B_{ij} (когда $n_i \neq 0$) — вопросами к базе A_i . Если $B_{ij} \subseteq A_i$, то называем вопрос B_{ij} к базе A_i уместным. Тавтологично ложную S -формулу $\textit{true} \rightarrow \textit{false}$ далее для краткости обозначаем $\perp$.

Пусть в (1) $F \neq \perp$, $n_i \neq 0$ и Φ_{ij} имеют вид

$$\Phi_{ij} = \bigvee_{k=1}^{l_{ij}} (C_{ijk} \wedge \Psi_{ijk}), \quad (2)$$

где $l_{ij} \neq 0$, иначе $\Phi_{ij} = false$. Предположим, что

$$\begin{aligned} \exists i^* \in 1, \dots, m, \quad n_{i^*} \neq 0, \\ j^* \in 1, \dots, n_{i^*}, \quad B_{i^*j^*} \subseteq A_{i^*}. \end{aligned} \quad (3)$$

т.е. вопрос $B_{i^*j^*}$ к базе A_{i^*} уместен.

Определение 2 (α -преобразование). Если S -формула F имеет вид (1)–(3), то α -преобразованием называется такое отображение $\alpha: L_S \rightarrow L_S$, что

$$\begin{aligned} \alpha(F) = true &\rightarrow \left(\bigvee_{i=1, i \neq i^*}^m \Omega_i^* \right) \vee \Omega_{i^*}^*, \\ \Omega_{i^*}^* &= \bigvee_{k=1}^{l_{i^*j^*}} \Omega_{i^*j^*k}^*, \\ \Omega_{i^*j^*k}^* &= \\ &= \left[\{A_{i^*} \cup C_{i^*j^*k}\} \wedge \Psi_{i^*j^*k} \wedge \left(\bigwedge_{j=1, j \neq j^*}^{n_{i^*}} (B_{i^*j} \rightarrow \Phi_{i^*j}) \right) \right], \end{aligned} \quad (4)$$

если $\Phi_{i^*j^*} \neq false$, иначе при $m > 1$ ГП $\Omega_{i^*j^*k}^*$ удаляются, а при $m = 1$ $\alpha(F) = \perp$.

Здесь и далее, когда это не вызывает коллизий, конъюнкты рассматриваются как множества входящих в эти конъюнкты элементов. В случае появления в конъюнкте $\{A_{i^*} \cup C_{i^*j^*k}\}$ из (4) контрарных литералов, формула $\alpha(F)$ упрощается заменой на $\perp$ при $m = l_{i^*j^*} = 1$, иначе удалением $\Omega_{i^*j^*k}^*$. Удаляются и дубли ГП при совпадении $\Omega_{i^*j^*k}^*$ с Ω_i , $i \neq i^*$.

Пусть $\alpha(F) \neq \perp$ и выполняется условие типа (3). Применяя α -преобразование к $\alpha(F)$, получим S -формулу $\alpha^2(F)$ и т.д. до тех пор, пока очередной результат еще отличен от $\perp$ и выполняется (3).

Пусть на некотором шаге r_1 ($r_1 \geq 0$) $\alpha^{r_1}(F) \neq \perp$, но α -преобразование уже не применимо ввиду невыполнения (3), т.е. $\forall i \in 1, \dots, m^1$, где m^1 – количество баз в $\alpha^{r_1}(F)$, база A_i будет:

1) либо висячей ($n_i = 0$) и, ввиду упомянутых упрощений, отличной от $false$, либо

2) невисячей ($n_i \geq 1$) и такой, что $\forall j \in 1, \dots, n_i$ вопросы B_{ij} к базе A_i неуместны (под выражениями A_i , B_{ij} понимаются новые конъюнкты в формуле $\alpha^{r_1}(F)$).

Определение 3 (β -преобразование). Пусть I_1, I_2 – множества индексов $i \in 1, \dots, m^1$, для

которых базы A_i удовлетворяют условиям 1) и 2) соответственно, и

$$\alpha^{r_1}(F) = true \rightarrow \bigvee_{i=1}^{m^1} \Omega_i^1,$$

где

$$\begin{aligned} \forall i \in I_1 \quad \Omega_i^1 &= A_i \quad \text{и} \\ \forall i \in I_2 \quad \Omega_i^1 &= A_i \wedge \left(\bigvee_{j=1}^{n_i} (B_{ij} \rightarrow \Phi_{ij}^1) \right). \end{aligned}$$

Назовем β -преобразованием такое отображение $\beta: L_S \rightarrow L_S$, что

$$\begin{aligned} \beta(\alpha^{r_1}(F)) &= true \rightarrow \bigvee_{i=1}^{m^1} \Omega_i^*, \\ \forall i \in 1, \dots, m^1 &= I_1 \cup I_2 \quad \Omega_i^* = \Omega_i \wedge P_i^1, \\ \forall i \in I_1 \quad P_i^1 &= (A_i \rightarrow false), \\ \forall i \in I_2 \quad P_i^1 &= \left(A_i \rightarrow \bigvee_{j=1}^{n_i} B_{ij} \right). \end{aligned}$$

Формирование условия

$$P^1 = P(\alpha^{r_1}(F)) = \bigwedge_{i=1}^{m^1} P_i^1$$

называем далее (α, β) -гипотезированием.

3. S -ИСЧИСЛЕНИЯ И ИХ СВОЙСТВА

На основе α - и β -преобразований формул языка L_S введем два исчисления.

Пусть $F^1 = \beta(\alpha^{r_1}(F))$. Применим к F^1 α -преобразование максимально возможное число раз r_2 ($r_2 \geq 1$). Если $\alpha^{r_2}(F^1) \neq \perp$, то применим β -преобразование, получив $F^2 = \beta(\alpha^{r_2}(F^1))$ и т.д. Пусть $F^n = \beta(\alpha^{r_n}(F^{n-1}))$ и впервые $\alpha^{r_{n+1}}(F^n) = \perp$. В результате n -кратного (α, β) -гипотезирования получили условие

$$P = \bigwedge_{i=1}^n P^i, \quad P^i = P(\alpha^{r_i}(F^{i-1})),$$

именуемое гипотезой-минорантой (ГМ).

Определение 4 (S -исчисления $\mathbb{C}_\alpha$, $\mathbb{C}_\beta$). S -исчисления $\mathbb{C}_\alpha = (\perp, \{\alpha\})$ и $\mathbb{C}_\beta = (\perp, \{\alpha, \beta\})$ – это исчисления в языке L_S с аксиомой $\perp$ и указанными преобразованиями в качестве правил вывода. Здесь выводы формулы F – это конечные последовательности формул, начинающиеся с F , заканчивающиеся формулой $\perp$ и с промежуточными формулами, получаемыми в $\mathbb{C}_\alpha$ с помощью только α -преобразований, а в $\mathbb{C}_\beta$ – с помощью преобразований α и β , как в построенной ранее для формулы F последовательности:

$F^0, F^1, F^2, \dots, F^{n+1}$, где $F^0 = F$. Вывод S -формулы F суть ее опровержение.

Теорема 1. Формула F^0 языка L_S противоречива тогда и только тогда, когда F^0 выводима в $\mathbb{C}_\alpha$.

Пусть $\{T, G, \Delta\} \subset L_S$, T – теория предметной области (контент некоторой базы знаний), G – наблюдаемый эффект. Объявление наблюдяемого признается гипотеза Δ , удовлетворяющая требованиям:

- а) из $T \wedge \Delta$ в исчислении $\mathbb{C}_\beta$ выводимо G и
- б) $T \wedge \Delta$ – непротиворечиво.

С использованием теоремы 1 доказуема следующая

Теорема 2. Пусть $F^0 = (T \wedge \neg G)^{L_S}$ – образ отрицания формулы $(T \rightarrow G)$ в языке L_S . Пусть также $P \in L_S$ и P построено (α, β) -гипотезированием в процессе вывода формулы F^0 в исчислении $\mathbb{C}_\beta$. Тогда $P \leftrightarrow (T \rightarrow G)$ и для любой гипотезы Δ , объясняющей G , справедливо $\Delta \rightarrow P$.

В разделе 4 демонстрируется переход от гипотез-минорант P к достаточным условиям путем удаления элементов, несовместимых с теорией T (примеры 1 и 2), или путем привлечения эмпирических знаний (пример 2). При этом получаются либо сами причины Δ эффекта G (достоверные или правдоподобные), либо релевантные обстоятельства (пример 1), из которых в рамках теории T средствами исчислений $\mathbb{C}_\alpha$ или $\mathbb{C}_\beta$ оказываются выводимыми и сами причины.

4. ПРИМЕРЫ ПРИМЕНЕНИЯ ИСЧИСЛЕНИЙ

Пример 1. Рассмотрим простой пример из области диагностики автомобиля (знания рассматриваемого ниже типа сформулированы в [1] в первом приближении для иллюстрации концепции экспертных систем в ИИ).

1) Пусть некоторые пропозициональные переменные языка L_S имеют следующий смысл: Φ – Фары горят; B – в Баке топливо есть; K – топливо поступает в Карбюратор; D – топливо поступает в Двигатель; V – двигатель Вращается; I – автомобиль Исправен; Z, A, C – проблемы, соответственно, в свечах Зажигания, в Аккумуляторе и в Стартере.

2) Пусть на фоне неисправности автомобиля выявлены следующие факты: Фары горят, двигатель вращается и топливо поступает в карбюратор.

3) Предположим, что, помимо этих фактов, базовая теория T включает знания: Если топливо поступает в двигатель и двигатель вращается, то проблема в свечах зажигания. Если в баке топливо

есть и топливо поступает в карбюратор, то топливо поступает в двигатель. Если двигатель не вращается, то при горящих фарах проблема в стартере, при негорящих – в аккумуляторе. Если автомобиль исправен, то проблем в свечах зажигания, стартере и аккумуляторе нет.

Базовая теория T представима следующей S -формулой вида

$$T = \{\Phi, B, B\} \wedge [(\{D, B\} \rightarrow Z) \wedge (\{B, K\} \rightarrow D) \wedge \wedge [\neg B \rightarrow (\text{true} \wedge ((\Phi \rightarrow C) \wedge (\neg \Phi \rightarrow A)))] \wedge \wedge (I \rightarrow \{\neg Z, \neg C, \neg A\})].$$

Наблюдаемое G состоит в неисправности автомобиля: $G = \neg I$, а вывод объяснения в S -исчислении $\mathbb{C}_\beta$ начинается с формулы $F^0 = (T \wedge \neg G)^{L_S}$. Ее база $\{\Phi, B, B, I\}$ при получении F^1 (однократным применением α -преобразования к F^0) дополняется элементами множества $\{\neg Z, \neg C, \neg A\}$.

Применением β -преобразования получается гипотеза-миноранта

$$P = P^1 = P(\alpha^1(F^0)) = P(F^1) = (\{\Phi, B, B, I, \neg Z, \neg C, \neg A\} \rightarrow (D \vee K \vee \neg B)).$$

После представления ее в форме элементарной дизъюнкции и удаления из нее членов, $\neg \Phi$, $\neg B$, $\neg B$, $\neg I$, контрарных элементам базы из F^1 , а также тривиального члена $Z \vee C \vee A$ получается объяснение $D \vee K$ в классе релевантных обстоятельств. Действительно, далее в рамках исчисления $\mathbb{C}_\alpha$ промежуточным элементом вывода формулы $(T \wedge \neg G \wedge (D \vee K))^{L_S}$, включившей полученное обстоятельство, оказывается факт Z как причина неисправности.

Пример 2. Рассмотрим другой иллюстративный пример из области медицины [1], для краткости не приводя детальной семантики переменных (статья [1] посвящена представлению знаний в диагностике с достоверными и эмпирическими знаниями).

Пусть e, h – симптомы заболевания. Известны болезни t, d, a . Пусть $\{t \rightarrow e, a \rightarrow h\}$ – множество достоверных причинно-следственных связей базовой теории T в терминах симптомов e, h и болезней t, a . Пусть $\{d \rightarrow h, a \rightarrow e\}$ – правдоподобные (недостоверные) причинно-следственные связи (ППСС) (эмпирические данные “прошлого опыта”). Пример интересен, в частности, и тем, что словарь базовой теории T не охватывает словарь эмпирических ППСС (нет болезни d). Тем не менее, из гипотезы-миноранты, выводимой средствами (α, β) -гипотезирования, получаются все логически возможные в рассматриваемом контексте диагнозы.

Если наблюдается симптом $G = e$, то

$$F^0 = (T \wedge \neg G)^{L_s} = \\ = (\text{true} \rightarrow \neg e \wedge [(t \rightarrow e) \wedge (a \rightarrow h)]).$$

В исчислении $\mathbb{C}_\beta$ при выводе F^0 формируется гипотеза-миноранта

$$P = \bigwedge_{i=1}^2 P(\alpha^i(F^{i-1})) = \\ = (\neg e \rightarrow (t \vee a)) \wedge (\neg e, a, h \rightarrow t).$$

При этом $r_1 = 0$, $r_2 = 3$, $r_3 = 2$ и $\alpha^3(F^2) = \perp$. Все элементарные дизъюнкции конъюнктивной нормальной формы (КНФ) $(e \vee t \vee a) \wedge (e \vee t \vee \neg a \vee \neg h)$ гипотезы-миноранты P следуют из причины t . По теореме 2 причина t объясняет симптом e .

Это объяснение не использует ППСС. С учетом же последних и того, что оба члена КНФ следуют из a и $a \rightarrow e$ соответственно, выводится еще одно (теперь только правдоподобное) объяснение: “причина – в болезни a , если верна ППСС $a \rightarrow e$ ”.

Пусть теперь наблюдаются одновременно два симптома: $G = e \wedge h$. Выявление причин использует вывод формулы

$$F^0 = (T \wedge \neg G)^{L_s} = \\ = (\text{true} \rightarrow \text{true} \wedge [(t \rightarrow e) \wedge (a \rightarrow h) \wedge (e \rightarrow \neg h)]).$$

При этом формируется гипотеза-миноранта

$$P = \bigwedge_{i=1}^3 P^i \leftrightarrow ((t \vee a \vee e) \wedge (\neg t \vee a \vee \neg e \vee h) \wedge \\ \wedge (t \vee \neg a \vee e \vee \neg h) \wedge (t \vee a \vee \neg e \vee h)) \quad (5)$$

и на ее основе три следующих объяснения. Два из них:

1) “комплекс из двух причин $\{t, a\}$ ” как достоверное объяснение,

2) “если верна ППСС $a \rightarrow e$, то причина в a ” (правдоподобное объяснение).

Третье объяснение получается из (5) использованием второй ППСС после расширения сигнатуры (5) переменной d путем замены каждой элементарной дизъюнкции на две других, получающихся добавлением в одну – переменной d , а во вторую – литералы $\neg d$. Каждая из элементарных дизъюнкций расширенной КНФ будет следовать из причин d , t и ППСС $d \rightarrow h$, что и явится третьим объяснением наблюдаемого комплекса симптомов $G = e \wedge h$.

СПИСОК ЛИТЕРАТУРЫ

1. Poole D. // J. Annals of Mathematics and Artificial Intelligence. 1994. V. 11. P. 33–50.
2. Kowalski R. / In: Computational Logic. Edition: In the History of Logic series. Eds D. Gabbay, J. Woods. Elsevier, 2014. P. 523–569.
3. Финн В.К. // Научно-техническая информация. Сер. 2. 2019. № 10. С. 1–34.
4. Васильев С.Н. // Информационные технологии и вычислительные системы. 2008. № 1. С. 3–19.
5. Kowalski R., Sadri F. // New Generation Computing. 2015. V. 33. P. 33–67.
6. Васильев С.Н., Жерлов А.К. // ДАН. 1995. Т. 343. № 5. С. 583–585.
7. Люгер Дж.Ф. Искусственный интеллект: стратегии и методы решения сложных проблем. М.: Изд. дом “Вильямс”, 2003. 864 с.

ABDUCTIVE REASONING IN EXPLANATION PROBLEMS OF THE OBSERVED

Academician of the RAS S. N. Vassilyev^a

^a Trapeznikov Institute of Control Sciences of the Russian Academy of Sciences, Moscow, Russian Federation

The problems of artificial intelligence, as well as control and decision-making with incomplete or inaccurate information, include a wide class of problems of abductive explanation, including tasks in terms of cause-effect. This paper is devoted to the logical formation of hypotheses that explain the observed effects. Means of representing knowledge and hypothesizing are proposed. A language is introduced that has the property of so-called substitutability. The properties of the language and the calculuses introduced on its basis provide a convenient combination of deduction and hypothesizing. Unlike the well-known logical methods of abduction, the proposed tools provide derivation of hypotheses (minorants) that are necessary and sufficient for a formal explanation of the observed. Based on the hypotheses-minorants, in combination with the basic theory of subject domain, reliable reasons for the observed are formed or relevant circumstances leading to these reasons are found. Moreover, in situations where there is also empirical data, these reasons and circumstances can also be formed in plausible versions. Examples from technology and medicine are considered.

Keywords: abductive reasoning, plausible inference, explanation of the observed, diagnosis, artificial intelligence, intelligent control