SEMINARIO DE LÓGICA II

Los límites de la lógica

Prof. Gustavo Fernández Díez

Curso 2009—2010
Primer Cuatrimestre

Universidad de Murcia

Índice abreviado

MÓDULO 0

Información académica

Ficha técnica

§ 0.1. Datos de la asignatura.

Nombre	Seminario de Lógica II (Los límites de la lógica)
Código	03L1
Créditos	6 créditos ECTS
Tipo	Optativa
Docencia	Impartida exclusivamente en el Campus Virtual SUMA
Duración	Cuatrimestral
Centro al que pertenece	Facultad de Filosofía
Titulación	Licenciado en Filosofía (Cursos 1º, 2º, 3º)
Otras titulaciones	Accesible por libre elección a todas las de 1º y 2º Ciclo

§ 0.2. Datos del profesor.

Profesor	Gustavo Fernández Díez
Categoría	Profesor Titular de Universidad
Centro	Facultad de Filosofía (Campus de Espinardo)
Departamento	Filosofía
Área de conocimiento	Lógica y Filosofía de la Ciencia
Despacho	3.65 (Facultad de Filosofía, Edif. Luis Vives, 3ª planta)
Teléfono	868 88 77 53 (hay contestador automático)
Correo electrónico	gfdezdp@um.es
Horario de atención presencial durante los períodos lectivos	Miércoles y jueves de 11:00 a 14:00, o bien mediante cita previa

§ 0.3. Presentación.

El contenido de esta asignatura corresponde a lo que podríamos llamar, "Lógica matemática avanzada". Contiene una parte introductoria (Módulos 1 y 2), en la que se desarrollan detalladamente los conceptos básicos de la lógica moderna. Y en el resto (Módulos 3 al 6), se emplean dichos conceptos en la exploración y prueba rigurosa de los principales resultados alcanzados, incluyendo los teoremas de incompletitud de Gödel.

Como asignatura ofertada por la Facultad de Filosofía, se hace especial hincapié en las implicaciones filosóficas de los resultados obtenidos.

§ 0.4. Conocimientos previos.

Aunque el desarrollo del temario no presupone conocimientos previos de la materia, el estilo expositivo puede resultar algo denso para los estudiantes de humanidades. Por tal razón se recomienda a este tipo de estudiantes haber cursado con anterioridad una asignatura introductoria, como la troncal de "Lógica" de la titulación de Filosofía.

Ello es una simple recomendación, no siendo requisito obligatorio haber cursado ninguna otra asignatura para poder matricularse de ésta.

Programa

§ 0.5. Objetivos.

Conceptuales: familiarizarse con los principales resultados metateóricos que muestran las limitaciones inherentes a la lógica formal; conocer el mecanismo de prueba de estos resultados de forma detallada y rigurosa; identificar sus implicaciones filosóficas más sobresalientes.

Procedimentales: reforzar la destreza en la utilización de lenguajes formales, así como en el manejo simultáneo de las perspectivas teórica y metateórica; reforzar la capacidad de razonamiento riguroso apoyado en un lenguaje simbólico.

Actitudinales: incrementar el interés por el estudio de los fundamentos de la matemática; incrementar el interés por la aplicación del razonamiento simbólico al análisis de problemas filosóficos.

§ 0.6. Programa de Teoría.

1. NOCIONES DE TEORÍA DE CONJUNTOS
— Introducción.
— Convenciones metalingüísticas previas.
— Nociones de teoría de conjuntos.
— La inducción matemática.

2. ELEMENTOS DE LÓGICA DE PRIMER ORDEN
— Términos.
— Fórmulas.
— Modelos.
— La consecuencia lógica.
— La sustitución.

— El cálculo de predicados.
— Teorías.

3. MODELOS NO ESTÁNDAR
— La aritmética de primer orden.
— Modelos no estándar de la aritmética.

4. NOCIONES DE TEORÍA DE LA RECURSIÓN
— Relaciones y funciones recursivas.
— El teorema de Matiyasevich.

5. EL TEOREMA DE TARSKI
— Aritmeticidad.
— El teorema de Tarski.
— Teorías recursivamente axiomatizables.

6. EL TEOREMA DE CHURCH Y LOS TEOREMAS DE GÖDEL
— La aritmética de Peano de primer orden.
— El teorema de Church.
— Los teoremas de Gödel.

§ 0.7. Programa de Prácticas.

Las Prácticas se repartirán en 14 Unidades, correspondientes a cada una de las semanas lectivas del Cuatrimestre a excepción de la primera, que no tendrá actividades asignadas.

La sección § 0.11 indica el procedimiento para el envío de las Prácticas, y en § 0.12 se hallará el *Cronograma*, con todas las actividades correspondientes a cada Unidad y sus plazos de entrega.

§ 0.8. Bibliografía básica.

Hunter, G., *Metalógica*, Madrid: Paraninfo, 1981.
Ladrière, J., *Limitaciones internas de los formalismos*, Madrid: Tecnos, 1969.
Machover, M., *Set Theory, Logic and their Limitations*, Cambridge: Cambridge University Press, 1996.
Zalabardo, J. L., *Introducción a la teoría de la lógica*, Madrid: Alianza, 2002.

Plan docente

§ 0.9. Metodología.

La metodología docente para esta asignatura se encuentra adaptada simultáneamente a la enseñanza no presencial y al Sistema Europeo de Créditos ECTS.

El principal material de trabajo será el presente Manual, que contiene íntegramente tanto la parte teórica de la asignatura como las Prácticas obligatorias diseñadas para su correcto seguimiento.

La impartición de la asignatura está distribuida en Unidades Didácticas Semanales, correspondientes a cada una de las semanas lectivas de que consta el Cuatrimestre. Al comienzo de cada semana el profesor hará una presentación de la Unidad, resaltando aquellas partes a las que los estudiantes deban prestar especial atención y recordando las actividades a realizar. A continuación se dará paso al Foro de debate, moderado por el profesor, para exponer comentarios y dificultades surgidos en relación con esa Unidad.

§ 0.10. Manejo del Manual del Curso.

Este Manual ha sido confeccionado con el editor de textos LaTeX 2_ε, y la versión electrónica del mismo ha sido obtenida mediante el programa PDFLaTeX. Hay que advertir que el buscador de Acrobat® Reader™ no reconoce correctamente los acentos del archivo. Por ello, para intentar localizar una palabra acentuada, se debe introducir en la utilidad de búsqueda algún fragmento relevante de la misma que no contenga acentos (por ejemplo "filosof" para buscar "filosofía").

El profesor recomienda que cada estudiante lleve el archivo a una fotocopiadora (en un CD o lápiz de memoria), y se haga una copia encuadernada. Ésta debe imprimirse "a doble cara", que es para lo que está diseñada la maquetación de las páginas. El coste aproximado es de 8 euros.

En cuanto a la estructura interna del Manual del Curso, está dividido en distintos Módulos. El Módulo 0 corresponde a la información académica, y los Módulos 1–6 corresponden a los contenidos de la asignatura. Cada Módulo va subdividido en Secciones cortas, marcadas con "§", y numeradas por relación al Módulo y a su lugar de orden dentro de éste.

Asimismo, los Módulos están distribuidos en varios Apartados de mediana extensión, como por ejemplo el Apartado dedicado al *Plan docente* en el que estamos ahora. Dichos Apartados no van numerados, ni afectan a la numeración correlativa de Secciones dentro del Módulo.

A continuación de los Módulos hay una *Bibliografía general*, que amplía sustancialmente la escueta *Bibliografía básica* de § 0.8. Por último, el Manual incluye también dos Índices de contenidos: el *Índice abreviado*, colocado al principio, y el *Índice general*, situado al final, en el que aparecen los números de página de todas las Secciones numeradas de los distintos Módulos.

§ 0.11. Uso de SUMA.

Manual del Curso.

El presente Manual está depositado en SUMA, dentro de la página web correspondiente a la asignatura, en *Contenidos — Documentos*.

Teoría.

El profesor presentará cada Unidad Didáctica Semanal mediante un mensaje en el *Tablón*, y a continuación dará paso al *Foro*, para debatir cuestiones relacionadas con esa Unidad. Ello incluye las cuestiones puramente académicas, que están comprendidas en la Unidad 1.

Además de participar en el *Foro*, los estudiantes pueden plantear sus dudas de forma individualizada, mediante la herramienta *Tutoría*, o en el horario de atención presencial del profesor indicado en § 0.2. También pueden consultar la sección de *Faqs* (Preguntas más frecuentes), que contiene la información más básica sobre el seguimiento de la asignatura.

Prácticas.

Las Prácticas se pueden entregar, bien en fichero informático a través de SUMA, o bien en papel, depositándolas en el casillero del profesor en Conserjería del Edificio Luis Vives.

Para depositar las Prácticas en SUMA se debe confeccionar un archivo titulado "Semana *X*" (donde "*X*" es el número de la semana correspondiente), y colgarlo a través de la ruta: *Contenidos — Mis Contenidos Alumnos — Subir archivo*. Dichos archivos se almacenarán automáticamente en una carpeta individualizada para cada estudiante, a la que sólo tiene acceso el profesor.

Para confeccionar el archivo informático con las soluciones a las Prácticas, se puede utilizar un editor de textos como MS Word® u otro, o bien escanear directamente una hoja manuscrita.

Los signos que no se encuentren en el editor de textos se pueden sustituir por otros a voluntad, señalando la correspondencia al principio de la práctica (por ejemplo: "utilizaré el signo «#» para las fórmulas de desigualdad"). Otra opción es reemplazar el signo correspondiente por su forma de lectura en voz alta, que viene siempre indicada en el Manual del Curso (poniendo, por ejemplo: "x distinto de y" en lugar de: "$x \neq y$").

El profesor contestará a todas las Prácticas de forma individualizada mediante una *Tutoría*. En el caso de las Prácticas marcadas con asterisco, colgará las Soluciones en *Contenidos — Documentos*, una vez finalizado el plazo de entrega.

§ 0.12. Cronograma.

Unidad didáctica		*Contenidos teóricos*	*Prácticas obligatorias*	*Plazo de entrega*
1	21 sep.—25 sep.	§ 0.1 —§ 0.19	———————	————
2	28 sep.—02 oct.	§ 1.1 —§ 1.14	§ 1.10; § 1.13; § 1.15	02/10/2009
3	05 oct.—09 oct.	§ 1.16—§ 1.31	§ 1.17; § 1.20; § 1.22; § 1.24; § 1.29	09/10/2009
4	13 oct.—16 oct.*	§ 1.32—§ 1.45	§ 1.40; § 1.42; § 1.46	16/10/2009
5	19 oct.—23 oct.*	§ 1.48—§ 2.11	§ 1.47; § 2.12	23/10/2009
6	26 oct.—30 oct.	§ 2.13—§ 2.25	§ 2.19; § 2.21; § 2.26	30/10/2000
7	02 nov.—06 nov.	§ 2.27—§ 2.51	§ 2.28; § 2.32; § 2.34; § 2.37	06/11/2009
8	09 nov.—13 nov.	§ 2.53—§ 2.76	§ 2.38; § 2.52; § 2.58; § 2.64	13/11/2009

Unidad didáctica		Contenidos teóricos	Prácticas obligatorias	Plazo de entrega
9	16 nov.—20 nov.	§ 2.78 —§ 3.18	§ 2.73; § 2.77; § 2.86	20/11/2009
10	23 nov.—27 nov.	§ 3.19 —§ 4.6	§ 3.4; § 3.10; § 3.29	27/11/2009
11	30 nov.—04 dic.	§ 4.8 —§ 5.8	§ 4.7; § 4.10; § 4.16; § 5.9	04/12/2009
12	09 dic.—11 dic.*	§ 5.10 —§ 5.24	§ 5.11; § 5.16; § 5.22; § 5.23	11/12/2009
13	14 dic.—18 dic.	§ 5.25 —§ 6.8	§ 5.33; § 5.39	18/12/2009
14	21 dic.—08 ene.†	§ 6.10 —§ 6.29	§ 6.3	08/01/2010
15	11 ene.—15 ene.	§ 6.31 —§ 6.41	§ 6.23; § 6.37	15/01/2010

* Semanas de menos de 5 días lectivos, por festividades varias: 12 de octubre, Fiesta Nacional de España; 22 de octubre, Bienvenida Universitaria; 7 de diciembre, lunes siguiente al Día de la Constitución; 8 de diciembre, Inmaculada Concepción.
† La semana 14ª va "partida" del siguiente modo: lunes-martes 21-22 de diciembre, justo antes de las vacaciones de Navidad; y jueves-viernes 7-8 de enero, a la vuelta de las mismas (total: 4 días lectivos).

§ 0.13. Dedicación estimada.

Las Unidades Didácticas Semanales han sido diseñadas para una dedicación estimada de 8 horas de estudio. Esas horas pueden distribuirse, a título orientativo, de la siguiente manera:

- 3 horas dedicadas a la lectura comprensiva de los Contenidos teóricos de la Unidad (unos 20 minutos por página, aproximadamente);
- 3 horas dedicadas a la realización de las Prácticas;
- 2 horas dedicadas la lectura del Tablón, así como a Tutorías y al Foro.

Se ha intentado que las distintas Unidades didácticas estén equilibradas en cuanto a la carga de trabajo que suponen. Para ello se ha tenido en cuenta no sólo el número de páginas del Manual del Curso que comprenden, sino también la densidad del contenido, y la cantidad y dificultad de las Prácticas asignadas. En las semanas de menos de 5 días lectivos, se ha efectuado una reducción proporcional.

El Real Decreto 1125/2003 del Sistema Europeo de Créditos, fija en 60 créditos la carga anual para los diferentes Planes de Estudio. Ello equivale a 5 asignaturas como ésta por Cuatrimestre. Suponiendo que todas exijan una dedicación similar (incluyendo las horas de asistencia a clase en las asignaturas presenciales), obtenemos un total de 40 horas semanales. Esto es, el estándar de jornada laboral a tiempo completo.

El estudiante que cumpla este horario de trabajo rigurosamente de lunes a viernes, no tiene por qué utilizar en absoluto los fines de semana, o los periodos vacacionales como Navidad y Semana Santa.

Por último, la preparación de la evaluación de la asignatura se estima en unas 30 horas de trabajo adicional. Éstas deben ubicarse preferentemente en el período de exámenes, una vez acabada la docencia efectiva.

Evaluación

§ 0.14. Fechas de examen (calendario provisional).*

Convocatoria	*Fecha*	*Hora*	*Edificio*	*Planta*	*Aula*
Febrero	02/02/2010	09:00	Luis Vives (Punto 12 del Campus)	–1	a.d.†
Junio	12/06/2010	09:00	"	"	"
Septiembre	07/09/2010	09:00	"	"	"

* El calendario oficial se anunciará previamente a cada convocatoria.

† A determinar.

§ 0.15. Evaluación de la Teoría.

1. La evaluación de la presente asignatura se basará en la presentación de un trabajo, más la realización de un examen escrito presencial de 2 horas de duración. Cada uno de esos dos elementos serán calificados independientemente de 0 a 10 puntos, y la nota obtenida en la asignatura será la media aritmética entre ambos.

2. En cada convocatoria, la fecha límite de entrega del trabajo coincidirá con el momento del examen. También se podrá presentar con anterioridad, por cualquiera de los procedimientos indicados en § 0.11 para la entrega de las Prácticas.

3. El trabajo consistirá en elaborar individualmente soluciones escritas a 5 de los 10 *Problemas propuestos* cuyo listado aparece en § 0.17. Las soluciones presentadas se valorarán sobre 2 puntos cada una.

4. El examen constará de 10 preguntas de respuesta breve, a elegir 5. Serán preguntas relacionadas con los contenidos del presente Manual del Curso, y se valorarán sobre 2 puntos cada una. En § 0.18 se encontrará un *Modelo de examen*, a fin de orientar sobre el tipo de cuestiones que se pueden plantear y su nivel de dificultad.

5. En la calificación tanto del trabajo como del examen se valorará, por este orden: la adecuación a lo preguntado, el conocimiento de la materia, la perspicacia demostrada y la claridad expositiva.

6. No se establece calificación mínima, ni en el trabajo ni en el examen, como requisito para superar la asignatura, siempre que la media aritmética sea de 5 puntos o mayor. Los alumnos que no superen la asignatura podrán solicitar la aplicación de la nota obtenida en cualquiera de las dos partes para sucesivas convocatorias.

§ 0.16. Evaluación de las Prácticas.

1. De acuerdo con el sistema de créditos ECTS, se establece como condición obligatoria para poder superar esta asignatura la realización en los plazos estipulados de un mínimo del 65 % de las Prácticas, lo cual equivale a 9 de las 14 Unidades semanales de asignación de tareas.

2. A pesar de su carácter obligatorio para poder superar la asignatura, la valoración individualizada de las Prácticas se transmite a cada estudiante a efectos meramente informativos, y no será utilizada como elemento de juicio para determinar la calificación.

3. Los estudiantes que por motivos justificados se vean imposibilitados de cumplir el requisito de entrega de las Prácticas, podrán solicitar una dispensa al Departamento de Filosofía (tel. 868 88 34 51). En caso de obtenerla no se les asignará tarea alternativa alguna, concurriendo directamente a la evaluación en igualdad de condiciones que el resto de sus compañeros.

§ 0.17. Problemas propuestos. (Véase § 0.15, Puntos 1 a 3.)

1. Mostrar que $\varnothing \subseteq A$ para cualquier conjunto A.

2. Siendo $\underline{\beta}$ y $\underline{\gamma}$ fórmulas cualesquiera de un lenguaje de primer orden, denotemos a una fórmula del tipo $\underline{(\beta \vee \gamma) \wedge \neg(\beta \wedge \gamma)}$ poniendo "$\underline{\beta \oplus \gamma}$". Especificar el valor de verdad de $\underline{\beta \oplus \gamma}$ bajo una evaluación ε, de modo similar a como se hace en § 2.38. Justificar la respuesta.

3. Sea $\mathscr{L}$ un lenguaje de primer orden carente de símbolos peculiares. Construir una oración $\underline{\alpha}$ que sea verdadera en un modelo para $\mathscr{L}$ si y sólo si su universo es un conjunto unitario.

4. Probar que si Φ es cualquier conjunto de oraciones de un lenguaje de primer orden, entonces
$$\mathit{teo}\,(\,\mathit{teo}\,(\Phi)\,) \quad = \quad \mathit{teo}\,(\Phi)$$

5. Sea $\mathscr{N}^\star$ cualquier modelo de Ω, y sea ι aquella correspondencia entre $\mathbb{N}$ y $\mathbb{N}^\star$ definida por
$$\iota(n) \;=\; \underline{s_n}^{\mathscr{N}^\star} \qquad\qquad \text{para todo } n \in \mathbb{N}$$
   (a) Probar que para cualesquiera $n, m \in \mathbb{N}$, si $n \neq m$, entonces $\iota(n) \neq \iota(m)$.
   (b) Probar que ι es una inmersión de $\mathscr{N}$ en $\mathscr{N}^\star$.
   (c) Probar que ι es *la única* inmersión posible de $\mathscr{N}$ en $\mathscr{N}^\star$.
   (d) Probar que $\mathscr{N}^\star$ es un modelo no estándar si y sólo si existe algún $a \in \mathbb{N}^\star$ tal que
$$a \neq \underline{s_n}^{\mathscr{N}^\star} \qquad\qquad \text{para todo } n \in \mathbb{N}$$

6. Sea $\mathscr{L}$ *cualquier* lenguaje de primer orden. Mostrar que si una oración de $\mathscr{L}$ es satisfecha por modelos de cualquier cardinalidad finita, también lo será por algún modelo infinito.

7. Sea R una relación numérica n-aria. Mostrar que si tanto R como su complementaria son recursivamente enumerables, entonces R es recursiva.

8. Sea Δ una teoría de $\mathscr{A}$ inconsistente, y sea $\underline{\alpha} \in \Upsilon_n$ para algún $n \in \mathbb{N}$.

    (a) Determinar las relaciones n-arias, si las hay, que $\underline{\alpha}$ representa débilmente en Δ.

    (b) Determinar las relaciones n-arias, si las hay, que $\underline{\alpha}$ representa fuertemente en Δ.

    Razonar la respuesta.

9. Sea $\mathscr{M}$ cualquier modelo para $\mathscr{A}$ cuyo universo conste únicamente de un elemento.

    (a) Probar que todas las oraciones de Π_0 son verdaderas en $\mathscr{M}$.

    (b) Probar que la oración $\underline{s_0 \neq s_1}$ *no* pertenece a Π_0.

10. Sea Δ aquella teoría de $\mathscr{A}$ generada por todos los axiomas de la forma

$$\underline{s_n + s_m = s_{n+m}}$$

    para cualesquiera números naturales n y m.

    (a) Indicar cómo se podría definir un procedimiento mecánico para decidir si un número natural cualquiera es el código de un axioma de Δ o no.

    (b) Indicar cómo se podría definir un procedimiento de enumeración para todos los códigos de oraciones de Δ.

    (c) Probar que Δ es indecidible, pero que el resultado de añadir a sus axiomas la oración $\underline{0 \neq 0}$, es una teoría decidible.

§ 0.18. Modelo de examen.

EXAMEN DEL SEMINARIO DE LÓGICA II

Universidad de Murcia

- Deberá contestar a **5** preguntas, de forma breve y concisa.
- La duración máxima de este examen es de 2 horas.
- Las Soluciones (así como las fechas de calificación y revisión) podrán consultarse una vez finalizado el examen, en SUMA, dentro del *Tablón* de la asignatura.

1. Decimos que dos conjuntos tienen "la misma cardinalidad" cuando sus elementos se pueden poner en correspondencia biunívoca. ¿En qué consiste eso?
2. (a) ¿Qué es el grado de una fórmula?
   (b) ¿Cuáles son las fórmulas de grado 0?
3. Especificar el valor de una fórmula condicional $\underline{\beta \to \gamma}$ bajo una evaluación ε, según los valores que reciban las fórmulas $\underline{\beta}$ y $\underline{\gamma}$ bajo esa evaluación.
4. Enunciar sucintamente las 5 condiciones para que ι constituya una inmersión de $\mathscr{N}$ en $\mathscr{N}^\star$.
5. Sea $\varepsilon^\star$ cualquier evaluación que satisfaga el conjunto de Skolem. Mostrar que $\underline{v_1}^{\varepsilon^\star} \neq \iota(n)$, para cualquier $n \in \mathbb{N}$ e inmersión ι de $\mathscr{N}$ en el modelo de $\varepsilon^\star$.
6. Demostrar que cualquier relación diofántica es recursivamente enumerable.
7. Enunciar las condiciones para que una fórmula $\underline{\alpha} \in \Upsilon_n$ represente una relación n-aria R en una teoría Δ:
   (a) débilmente;
   (b) fuertemente.
8. Enunciar el teorema de Tarski. Explicar sucintamente por qué dicho teorema implica que $\mathscr{A}$ no puede funcionar como su propio metalenguaje.
9. Sea Δ cualquier teoría de $\mathscr{A}$ tal que $\Delta \cup \Pi_2$ es consistente, y sea $\underline{\pi}$ la conjunción de los nueve axiomas de Π_2. Mostrar que para cualquier oración $\underline{\alpha}$,
$$\underline{\alpha} \in teo\,(\Delta \cup \Pi_2) \qquad \text{si y sólo si} \qquad \underline{\pi \to \alpha} \in \Delta$$
10. Enunciar el primer teorema de incompletitud de Gödel, en sus dos versiones.

§ 0.19. Soluciones al Modelo de examen.

EXAMEN DEL SEMINARIO DE LÓGICA II

Soluciones orientativas

(Entre paréntesis se indican las páginas del Manual del Curso donde se pueden consultar más ampliamente las respuestas.)

1. Es un emparejamiento en el cual cada elemento de un conjunto tiene una única y exclusiva pareja en el otro, distinta de la que tienen los demás, y sin que quede ningún elemento por emparejar en ninguno de los dos conjuntos (p. 37).

2. (a) El número total de ocurrencias de conectivas y del símbolo de cuantificación universal en esa fórmula.
   (b) Las fórmulas atómicas (p. 52).

3.
$$\underline{\beta \to \gamma}^{\,\varepsilon} = \begin{cases} \mathbb{F} & \text{si } \underline{\beta}^{\,\varepsilon} = \mathbb{V} \text{ y } \underline{\gamma}^{\,\varepsilon} = \mathbb{F} \\ \\ \mathbb{V} & \text{en cualquier otro caso} \end{cases}$$
(p. 59).

4. 1. ι es una correspondencia biunívoca entre $\mathbb{N}$ y algún subconjunto de $\mathbb{N}^\star$;
   2. $\iota(0) = 0^\star$;
   3. para todo $n \in \mathbb{N}$, $\iota\,(\,s(n)\,) \;=\; s^\star\,(\,\iota(n)\,)$;
   4. para cualesquiera $n, m \in \mathbb{N}$, $\iota\,(n+m) \;=\; \iota(n) \;+^\star\; \iota(m)$;
   5. para cualesquiera $n, m \in \mathbb{N}$, $\iota\,(n \cdot m) \;=\; \iota(n) \;\times^\star\; \iota(m)$ (p. 82).

5. Como $\varepsilon^\star$ satisface Σ, hará verdaderas a todas las fórmulas de la forma $\underline{v_1 \neq s_n}$ para cada numeral $\underline{s_n}$. Lo cual implica
$$\underline{v_1}^{\,\varepsilon^\star} \neq \underline{s_n}^{\,\varepsilon^\star} \qquad\qquad \text{para todo } n \in \mathbb{N}$$
Ahora bien. Siendo $\mathscr{N}^\star$ el modelo de $\varepsilon^\star$, tenemos:
$$\begin{aligned} \underline{s_n}^{\,\varepsilon^\star} &= \underline{s_n}^{\,\mathscr{N}^\star} && \text{(al ser } \underline{s_n} \text{ un término cerrado)} \\ &= \iota\,(\underline{s_n}^{\,\mathscr{N}}) && \text{(al ser } \iota \text{ una inmersión de } \mathscr{N} \text{ en } \mathscr{N}^\star) \\ &= \iota\,(n) && \text{(por el valor de los numerales en } \mathscr{N}) \end{aligned}$$
Con lo cual queda demostrado $\underline{v_1}^{\,\varepsilon^\star} \neq \iota\,(n)$, para todo $n \in \mathbb{N}$ (p. 85).

6. Dada cualquier relación diofántica n-aria Q, por definición ha de existir una relación elemental R, tal que para cualquier $\vec{a} \in \mathbb{N}^n$

$$Q(\vec{a}) \qquad \text{si y sólo si} \qquad \text{existe una secuencia } \vec{b} \in \mathbb{N}^m \text{ tal que } R(\vec{a}, \vec{b})$$

   Entonces, puesto que R es elemental, y por tanto recursiva y r.e., basta con modificar un procedimiento de enumeración para R, de tal forma que borre los m últimos componentes de cada secuencia que genere, y vaya omitiendo repeticiones según ocurran. El resultado es un procedimiento de enumeración para Q (p. 95).

7. (a) Cuando para cualquier $\vec{a} \in \mathbb{N}^n$,

$$\begin{cases} \text{si} & R(\vec{a}) \text{ entonces } \underline{\alpha[s_{\vec{a}}]} \in \Delta \\ \text{si} & \sim R(\vec{a}) \text{ entonces } \underline{\alpha[s_{\vec{a}}]} \notin \Delta \end{cases}$$

   (b) Cuando para cualquier $\vec{a} \in \mathbb{N}^n$,

$$\begin{cases} \text{si} & R(\vec{a}) \text{ entonces } \underline{\alpha[s_{\vec{a}}]} \in \Delta \\ \text{si} & \sim R(\vec{a}) \text{ entonces } \underline{\neg\alpha[s_{\vec{a}}]} \in \Delta \end{cases}$$

   (p. 97).

8. Este teorema dice que la propiedad de ser el código de una oración verdadera en $\mathscr{N}$ no es aritmética. Por consiguiente, la noción de *ser verdadero en la interpretación estándar* no es representable mediante una fórmula de $\mathscr{A}$. Y ello implica que $\mathscr{A}$ no puede funcionar como su propio metalenguaje (pp. 104, 105).

9.

$$\begin{array}{llllll} \underline{\alpha} \in teo\,(\Delta \cup \Pi_2) & \text{si y sólo si} & \Delta \cup \{\underline{\pi}\} & \vdash & \underline{\alpha} \\ & \text{''} & \Delta \cup \{\underline{\pi}\} & \models & \underline{\alpha} \\ & \text{''} & \Delta & \models & \underline{\pi \rightarrow \alpha} \\ & \text{''} & & & \underline{\pi \rightarrow \alpha} \in \Delta \\ & & & & \text{(al ser } \Delta \text{ una teoría)} \end{array}$$

   (p. 116).

10. Versión semántica: para cualquier teoría de $\mathscr{A}$ que sea recursivamente axiomatizable y correcta con respecto a $\mathscr{N}$, existe alguna oración verdadera en $\mathscr{N}$ que no pertenece a dicha teoría (p. 119). Versión sintáctica: si Δ es una teoría de $\mathscr{A}$ que es consistente, recursivamente axiomatizable e incluye a Π_1, entonces existe una oración tal que ni ella ni su negación pertenecen a Δ (p. 121).

MÓDULO 1

Nociones de teoría de conjuntos

Introducción

§ 1.1. Observación (Objeto de la lógica). El *razonamiento deductivo* es aquel tipo de razonamiento en el cual la verdad de las *premisas* desde las que se razona implica la verdad de la *conclusión* a la que se llega. El ejemplo por antonomasia es

"Todos los hombres son mortales
Sócrates es un hombre
Por lo tanto:
Sócrates es mortal"

Este tipo de razonamiento se opone al llamado *"razonamiento inductivo"* (o *"razonamiento probable"*), en el cual la verdad de las premisas hace más verosímil o probable la verdad de la conclusión, pero no la implica, es decir: cabe la posibilidad de que las premisas sean verdaderas y la conclusión falsa. Un ejemplo de este otro tipo de razonamiento es:

"Casi todos los hombres mueren antes cumplir 100 años
Arturo Pérez-Reverte es un hombre
Por lo tanto:
Arturo Pérez-Reverte morirá antes de cumplir 100 años"

Como se puede apreciar, aún constatando que las dos premisas son verdaderas, cabe la posibilidad de que la conclusión sea falsa, es decir: cabe la posibilidad —aunque pequeña— de que Arturo Pérez-Reverte viva más de 100 años.

Pues bien: la *lógica* es la ciencia que tiene como objeto el estudio de las condiciones de validez del razonamiento deductivo. Es decir: el estudio de las condiciones bajo las cuales podemos considerar que un razonamiento es válido o inválido, correcto o incorrecto, desde el punto de vista del razonamiento deductivo.

§ 1.2. Observación (La lógica de primer orden clásica). En la actualidad, la teoría mejor y más refinada que poseemos sobre la esencia del razonamiento deductivo correcto es la disciplina conocida como *"lógica de primer orden clásica"*. Su descubrimiento fue obra del matemático y filósofo alemán Gottlob Frege y otros pensadores a finales del siglo XIX, y supuso un avance gigantesco en esta materia, sólo comparable al trabajo pionero que hizo, en la Antigua Grecia, Aristóteles. Es por ello que actualmente se utilizan las expresiones "lógica moderna", "lógica formal", "lógica simbólica" y "lógica matemática" —expresiones, por lo demás, todas ellas esencialmente sinónimas—.

Además de la lógica de primer orden clásica, hay otras teorías sobre el razonamiento deductivo correcto, que circundan a esta teoría y que algún día pueden contribuir a su mejora. Son las llamadas *"lógicas no clásicas"*, y están divididas en dos grandes grupos: las *lógicas extendidas* y las *lógicas alternativas*. Cada una de ellas propone diversas reformas a la lógica de primer orden clásica, casi siempre con el objeto de dar un tratamiento especial a determinados contextos de razonamiento. Por ejemplo, entre las lógicas extendidas, la llamada *"lógica modal"* trata de aquellos razonamientos deductivos en los que los conceptos de *necesidad* y *posibilidad* juegan un papel esencial.

La diferencia está en que las lógicas extendidas proponen meras ampliaciones de la lógica clásica, mientras que las lógicas alternativas proponen revisiones profundas, que de ser aceptadas supondrían el rechazo a la lógica de primer orden clásica tal y como está configurada en la actualidad.

Por lo demás, y sin minusvalorar la importancia en conjunto de las lógicas no clásicas, hay que señalar que hoy por hoy, siempre que se dice "lógica de primer orden" sin más, se sobrentiende que se trata de la lógica de primer orden *clásica*.

§ 1.3. Observación (Las ramas de la lógica). Para llevar a cabo su estudio, la lógica recurre a la utilización de los llamados *"lenguajes formales"*. Estos son lenguajes artificiales, especialmente diseñados para investigar en ellos las condiciones del razonamiento deductivo correcto, por así decirlo, "en bruto".

Una vez definido un lenguaje formal, se plantean inmediatamente dos tareas principales: una es la descripción de aquellas posibles *interpretaciones* que puedan convenir a los enunciados formulados en tal lenguaje, lo cual da lugar a lo que llamamos *"teoría de modelos"*. Y la otra es la composición de un particular conjunto de *reglas* que reproduzcan, en ese lenguaje, las reglas correspondientes al razonamiento deductivo correcto, a lo cual denominamos *"teoría de la prueba"* (o *"teoría de la demostración"*). Así, cualquier sistema de lógica, incluyendo las lógicas no clásicas, lleva asociadas una teoría de modelos y una teoría de la prueba características.

La teoría de modelos correspondiente a la lógica de primer orden clásica se ha desarrollado y crecido de una forma tan espectacular, que hoy día constituye por sí misma toda una ciencia autónoma, cuyos objetivos y resultados van mucho más allá de la mera investigación de las condiciones del razonamiento correcto. La teoría de la prueba, por su parte, también está muy desarrollada, pero la importancia de sus resultados es algo menor y más restringida.

Junto a estas disciplinas hay que situar a otras dos, la *teoría de conjuntos* y la *teoría de la recursión*, muy ligadas a la lógica por su historia y por su naturaleza, aunque estrictamente hablando persiguen objetivos distintos. La teoría de conjuntos consiste en el estudio de la noción de *conjunto* y la relación de *pertenencia*. Una tarea aparentemente tan sencilla reviste gran importancia, porque se ha comprobado que la práctica totalidad de los objetos matemáticos pueden ser representados en términos de conjuntos y combinaciones de conjuntos, y análogamente, que la práctica totalidad de los teoremas matemáticos pueden ser probados a partir de las leyes o axiomas de la teoría de conjuntos.

Este hecho convierte a la teoría de conjuntos, junto a la propia lógica, en uno de los principales pilares para el estudio de la fundamentación de la matemática. Y en general, las relaciones entre las dos disciplinas son muy variadas y profundas.

Por su parte, la teoría de la recursión (también llamada *"teoría de la computabilidad"*), se encarga del estudio de aquellos cálculos y manipulaciones simbólicas que tienen un carácter puramente *mecánico*, es decir: rígido y determinista. Esta es una tarea en sí misma independiente del estudio del razonamiento correcto, pero está ligada a ella, entre otras razones, por cuestiones históricas, ya que de hecho el principal motivo de que se comenzase a estudiar seriamente las manipulaciones mecánicas, fue el interés por verificar si las reglas del razonamiento deductivo estaban entre ellas.

Hoy día la teoría de la recursión está también incomparablemente desarrollada, y admite por supuesto un estudio del todo independiente de la lógica propiamente dicha. Bien al contrario, es una disciplina necesaria para desarrollar cualquier sistema de lógica de forma rigurosa y hasta sus últimas consecuencias.

§ 1.4. Observación (Panorama de la asignatura). En la presente asignatura tocaremos ligeramente cada una de las ramas que se acaban de mencionar. En concreto, el Módulo 4 de este Manual está dedicado a una breve presentación de la teoría de la recursión. También se encontrará, más adelante en el Módulo actual, un apartado sobre nociones básicas de teoría de conjuntos. Y el Módulo 2 dedica sus dos primeros apartados a definir nuestros lenguajes formales, y el resto a exponer los rudimentos de la teoría de modelos, incluyendo también una mínima incursión en la teoría de la prueba, dentro del apartado titulado *"El cálculo de predicados"*. No obstante, no nos engañemos, todo esto se hace aquí a un nivel mínimo de profundidad, y en un formato sumamente facilitado.

El verdadero objetivo de esta asignatura no es servir como introducción sistemática a las distintas ramas de la lógica moderna, sino exponer monográficamente los principales resultados obtenidos por ella. Y sucede que, de hecho, la mayor parte de estos resultados son *limitativos*, es decir: muestran las limitaciones de los métodos de la lógica formal para realizar determinadas tareas. De ahí el título de la asignatura.

Estos resultados se refieren principalmente a cómo se aplica la lógica de primer orden a cierta teoría matemática, la *aritmética*, que es quizás la más elemental y universal de las teorías matemáticas. La aritmética es sencillamente la parte de las matemáticas que estudia la serie de los números naturales: 0, 1, 2, 3, ..., y las cuatro operaciones básicas entre ellos, suma, resta, multiplicación y división. *A priori* parece que tendría que ser muy fácil someter a formalización lógica una teoría tan sencilla, que hasta los niños son capaces de manejar. Pero la respuesta dista mucho de ser la esperada, como ya iremos viendo.

Aunque nuestra asignatura trate, en definitiva, de teoremas negativos o limitativos, no por ello hay que pensar que carecen de importancia. Muy al contrario, se trata de resultados que por su gran dificultad, por su profundidad e interés, se cuentan entre los principales avances de la ciencia en su conjunto durante el último siglo. Además, poseen implicaciones filosóficas de largo alcance, sobre la naturaleza de la lógica y de las matemáticas, y hay quien ha planteado otras más lejanas, sobre propiedades fundamentales de la mente humana, por ejemplo, de todas las cuales hablaremos en su momento.

§ 1.5. Observación (Definiciones, lemas y teoremas). Entre las Secciones de este Manual se encontrarán: *Definiciones*, en las que se caracterizan los distintos objetos de estudio; *Observaciones*, en las que se hacen comentarios diversos y se perfilan mejor esos objetos; *Convenciones*, en las que se introducen normas de uso para referirnos a ellos; y *Lemas*, *Teoremas* y *Corolarios*, en los que se establecen los hechos más importantes que dichos objetos protagonizan. En cuanto a estos tres últimos, el *lema* suele presentar un resultado algo menor, que se emplea como apoyatura para demostrar posteriormente un *teorema*; mientras que el *corolario* es una consecuencia inmediata de algún teorema anterior.

Se comprobará que se han omitido las pruebas de algunos lemas y teoremas. Esto se ha hecho al objeto de centrar la discusión en su tema principal, y mantener la asignatura en unas dimensiones asequibles, que permitan desarrollarla por completo en un cuatrimestre. El estudiante interesado puede encontrar referencias de dichas pruebas en la *Bibliografía básica* de § 0.8, y en especial en el libro de Machover que allí se cita.

§ 1.6. Observación (Agradecimiento). El presente Manual se ha beneficiado de las observaciones y sugerencias de las distintas promociones de alumnos que han cursado conmigo esta asignatura desde 1995, y especialmente de la profesora Bermejo Luque, que colaboró en su impartición durante el curso 2005–2006.

Convenciones metalingüísticas previas

§ 1.7. Observación (Lenguaje objeto y metalenguaje). Así como el físico que estudia la ley de la gravedad está sometido a ella, y es ella la que lo mantiene pegado al suelo, y así como el médico que analiza la respiración no puede dejar de practicarla, del mismo modo nosotros, que vamos a dedicarnos a escudriñar ciertas formas de razonamiento, no tendremos más remedio que seguir razonando y utilizando muchos de los patrones de inferencia que constituyan nuestros objetos de estudio.

Es por ello que debemos prestar atención. En esta sección preliminar vamos a perfilar aspectos de nuestra forma de expresarnos y de razonar a lo largo de la presente asignatura. Se trata, esencialmente, de adoptar convenciones y ponernos de acuerdo sobre cómo *nosotros* vamos a usar el lenguaje y a entendernos de aquí en adelante. Convenciones, por lo demás, que son absolutamente necesarias y de uso generalizado para cualquier estudio moderno de la lógica en serio.

Vamos a hablar en fin, de lo que técnicamente se denomina el *"metalenguaje"*. Nuestro metalenguaje no será más que el castellano vulgar y corriente, aunque enriquecido con algunos símbolos y términos técnicos, y algo modificado por las precisiones que se van a hacer a continuación.

Por otra parte, a lo largo de este Manual —a partir del Módulo siguiente— definiremos diversos *lenguajes formales*, lenguajes totalmente artificiales y extraños. Dichos lenguajes actuarán como nuestro *lenguaje objeto*, es decir, como lenguajes que nosotros tendremos como objetos de estudio. Con frecuencia el lenguaje objeto escogido reflejará aspectos de nuestro metalenguaje, e incluso reflejará algunas de las convenciones que vamos a establecer nosotros aquí. Pero tenemos que recordar que son distintos, y que lo que valga para el lenguaje objeto no tiene por qué coincidir siempre —y no siempre coincidirá— con lo que vale para el nuestro.

§ 1.8. Observación (Letras y tipos de letra). En la presente asignatura utilizaremos para determinados propósitos letras caligráficas mayúsculas, concretamente

$$\mathscr{A} \qquad \mathscr{K} \qquad \mathscr{L} \qquad \mathscr{M} \qquad \mathscr{N} \qquad \mathscr{R} \qquad \mathscr{S}\ \ (\textit{ese}) \qquad \mathscr{T}\ \ (\textit{te})$$

así como tres letras de las llamadas *"de doble trazo"* (o *"negrita de pizarra"*), en particular

$$\mathbb{F} \qquad \mathbb{N} \qquad \mathbb{V}$$

También utilizaremos algunas letras del alfabeto griego, concretamente las mayúsculas

Γ (*Gamma*) Δ (*Delta*) Λ (*Lambda*) Π (*Pi*) Σ (*Sigma*)

Υ (*Ípsilon*) Φ (*Phi*, leído *"fi"*) Ψ (*Psi*) Ω (*Omega*)

y las minúsculas

α (*alfa*) β (*beta*) γ (*gamma*) δ (*delta*)

ε (*épsilon*) ι (*iota*) π (*pi*)

Recurriremos a veces al *subíndice*, como en "a_2" (leído *"a sub 2"*), y al *superíndice*, como en "a^4" (leído *"a súper 4"*, excepto un determinado uso en el presente Manual, que indicaremos en su momento). Utilizaremos "$=$" con el significado *es igual a*; "$\geqslant$" como *es mayor o igual que*; y "$>$" como *es estrictamente mayor que.*

Utilizaremos también la abreviatura "cf.", que significa *confrontar* (equivalente al *"véase"*). Utilizaremos el asterisco o *estrella* "$\star$"; el apóstrofo "$'$" (que se lee *"prima"*); y la barra derecha "/" (que se puede leer sencillamente como *"barra"*). El signo de párrafo "§", como ya dijimos, se usa en este Manual para marcar la numeración de las Secciones (por lo que se leerá *"sección"*). Y en fin, otros signos especiales se definirán según se vayan introduciendo.

§ 1.9. Convención (Predicaciones impropias). Una *predicación impropia* es aquella en la cual se atribuye a un objeto, una propiedad que no tiene sentido plantearse del objeto en cuestión. Por ejemplo, si decimos

"El planeta Marte es juez del Tribunal Supremo"

estamos efectuando una predicación impropia, porque son las personas y no los planetas, las que pueden o no ser jueces, y en particular jueces del Tribunal Supremo.

A veces las predicaciones impropias contienen un significado metafórico. Por ejemplo en

"Italia es una bota que está a punto de darle una patada a Sicilia"

se quiere decir que el trazado del mapa de Italia recuerda a una bota que tiene a la isla de Sicilia delante de la puntera. En su significado literal dicho enunciado es no obstante falso, ya que Italia es un país, y no una bota, y como tal país no tiene capacidad —estrictamente hablando— de darle una patada a nada.

Pues bien: nosotros, a partir de este momento y ya a todo lo largo del Manual, adoptaremos la siguiente convención: siempre que nos encontremos o tengamos que tratar una predicación impropia, la tomaremos en su significado literal, y por lo tanto la consideraremos *falsa* a todos los efectos. Así, consideraremos que es falso, sin más, que Italia sea una bota, que Marte sea juez del Tribunal Supremo, y análogamente en todos los casos similares.

Lo mismo se aplica también a la predicación de relaciones entre distintos objetos, cuando se trata de una predicación impropia. Por ejemplo,

"Los planetas Júpiter y Saturno son primos hermanos"

atribuye una relación a estos dos planetas, la de ser primos hermanos. La diferencia con la atribución de propiedades es que aquí son varios los objetos involucrados, y lo que se predica es una cierta conexión entre ellos. Pero igualmente se trata de una predicación impropia, que tomada en su literalidad resulta falsa, y que nosotros consideraremos por consiguiente *falsa* a todos los efectos.

Aunque pueda parecer raro, durante esta asignatura vamos a encontrarnos con predicaciones impropias con cierta frecuencia, por lo que debemos prestar atención para seguir cuidadosamente la convención que acabamos de formular.

Por otra parte, en las contadas ocasiones en que queramos expresar un pensamiento metafórico, tendremos que advertirlo de forma clara, con expresiones como *"por hacer una comparación"*, *"por decirlo de alguna manera"* u otras similares, para subrayar que sólo en esos casos, y de manera excepcional, podemos atender a connotaciones del significado que excedan del puro significado literal.

§ 1.10. Práctica (Predicaciones impropias). Proponer un nuevo ejemplo de predicación propia y otro de impropia, relativas a propiedades. Repetir el ejercicio, pero esta vez con atribuciones de relaciones entre objetos. Especificar, de acuerdo con la convención precedente, cuál es el valor de verdad de los 4 ejemplos propuestos, esto es: determinar si cada uno de ellos es verdadero o falso.

§ 1.11. Convención (Uso de la negación). A partir de este momento y ya a todo lo largo del Manual, consideraremos que el valor de verdad de un enunciado negativo es el contrario al valor de verdad que tenga el enunciado negado. Es decir: si un enunciado es verdadero, entonces su negación será falsa, y viceversa.

Esto se aplica en particular a aquellos enunciados que contengan negaciones de predicaciones impropias. Así por ejemplo,

"El planeta Marte no es juez del Tribunal Supremo"
"Italia no es una bota que está a punto de darle una patada a Sicilia"
"Los planetas Júpiter y Saturno no son primos hermanos"

son los tres *verdaderos*, simplemente porque consisten en negaciones de enunciados que ya hemos dictaminado como falsos.

Esta convención sobre el uso de la negación puede parecer una simpleza, pero habremos de llevar cuidado si no queremos que nos juegue malas pasadas. Además, hay que tener presente que no sólo se aplica a las negaciones de las predicaciones impropias, sino a todos los enunciados negativos sin excepción, lo cual incluye a los restantes tipos de enunciados sobre los que vamos a establecer convenciones a continuación.

§ 1.12. Convención (Uso inclusivo de la disyunción). A partir de este momento y ya a todo lo largo del presente Manual de Curso, consideraremos que un enunciado disyuntivo es verdadero siempre que sea verdadero al menos uno de los disyuntos, es decir, al menos una de las opciones planteadas por la disyunción. Ello engloba los casos en que sólo uno de los disyuntos es verdadero, así como los casos en que son verdaderos dos o incluso más disyuntos, si es una disyunción triple, o cuádruple, etc. En todos estos casos, al enunciado disyuntivo lo consideraremos verdadero.

A este tipo de disyunción se le denomina técnicamente *"disyunción inclusiva"*. Nuestra convención equivale pues, a anunciar, que nuestro uso de la disyunción en la presente asignatura será conforme a la disyunción inclusiva que se acaba de describir.

Empecemos analizando el ejemplo

"O se mantienen las medidas de protección o se extinguirá el lince ibérico"

Este enunciado es claramente verdadero —por desgracia— porque en los próximos años, o bien sucede que se mantienen firmemente las medidas extraordinarias de protección de esta especie, o bien se extinguirá sin más remedio. Una de estas dos opciones es verdadera, no sabemos cuál —esperemos que sea la primera. Y debido a que una de las opciones es verdadera, la disyunción resulta a su vez verdadera.

Por su parte, en el siguiente ejemplo:

"O el Sol sale por el Este o se pone por el Oeste"

tenemos que las dos opciones propuestas son verdaderas. Esto normalmente nos dejaría perplejos, pero atendiendo a la convención que acabamos de efectuar, no deja lugar a dudas: se trata de una disyunción también *verdadera*. Y lo mismo pasa con

"O la nieve es blanca o la Tierra se mueve"

por idénticas razones, aunque en este caso los dos disyuntos propuestos no tengan nada que ver el uno con el otro.

En cuanto a las disyunciones triples, por su parte, basta con que uno de los disyuntos sea verdadero, pero también pueden serlo dos, o serlo los tres. Y algo análogo ocurre con las disyunciones cuádruples y mayores.

Sólo hay un caso, en definitiva, en que una disyunción resulta falsa, bajo nuestra convención. A saber: aquel en que son falsos todos y cada uno de los disyuntos que contenga.

§ 1.13. Práctica (Uso inclusivo de la disyunción). Poner un nuevo ejemplo de disyunción en la cual los disyuntos no tengan nada que ver el uno con el otro, y de forma que uno de los disyuntos sea verdadero y el otro falso. Poner otro ejemplo similar, pero con los dos disyuntos verdaderos. Finalmente, poner un último ejemplo con los dos disyuntos falsos, y dar valor de verdad a los 3 ejemplos propuestos (esto es, indicar si cada uno de ellos es verdadero o falso).

§ 1.14. Convención (Uso del condicional material). Los enunciados condicionales son aquellos en los que se antepone un hecho como condición de otro:

"Si no operamos inmediatamente, esta persona morirá"

En estos casos llamamos *"antecedente"* a la condición que se antepone (en nuestro ejemplo, el no operar inmediatamente), y *"consecuente"* a lo que ocurre como consecuencia de esa condición (la persona morirá).

Los enunciados condicionales se pueden parafrasear de diferentes formas, pero la idea de que hay una cosa que se pone como condición de otra, es fácilmente detectable:

"Si un hombre va armado, es peligroso"
"Si un hombre va armado, entonces es peligroso"
"Un hombre es peligroso si va armado"
"Un hombre es peligroso cuando va armado"
"Cuando un hombre va armado, es peligroso"
"Siempre que un hombre va armado, es peligroso"
Etc.

Todos ellos son enunciados condicionales, y todos ellos son, evidentemente, sinónimos.

Pues bien: nosotros, a partir de este momento y ya a todo lo largo del Manual, vamos a adoptar una convención muy singular con respecto a los enunciados condicionales: vamos a considerar a todos los enunciados condicionales verdaderos, excepto en el caso en que el antecedente sea de hecho verdadero y el consecuente falso. Ello significa que un enunciado condicional será falso cuando tenga antecedente verdadero y consecuente falso, y será verdadero en cualquier otro caso, y con total independencia de cualquier otra consideración.

A este tipo de condicional se le llama *"condicional material"*, y se dice de él que es *"veritativo–funcional"*, porque su valor de verdad está en función exclusivamente de los valores de verdad de sus componentes. Y no depende de otras cosas, como la conexión causal entre ellos, o la relevancia que pueda tener el uno para el otro.

Así por ejemplo

"Si Madrid está en Francia, entonces París está en Australia" (1)

es *verdadero*, aunque no tenga nada que ver una cosa con la otra, simplemente porque de hecho "Madrid está en Francia" es falso, y "París está en Australia" es también falso, es decir: porque antecedente y consecuente son, en este caso, ambos falsos.

Aunque el enunciado (1) pueda parecer un disparate, lo cierto es que no se encuentra bajo la única circunstancia posible, de acuerdo con nuestra convención, para que un condicional sea falso, a saber: que antecedente sea verdadero y consecuente falso. Y es por ello que, no siendo falso, no le queda más remedio que ser verdadero.

Por similares razones son también verdaderos los siguientes:

"Si París está en Francia, entonces Madrid está en España"
"Si Madrid está en Francia, entonces Madrid está en España" (!)

y *falso*, sin embargo —esta vez sí— un enunciado como:

"Si Madrid está en España, entonces París está en Australia"

El tipo de condicional que estamos postulando aquí es ciertamente idiosincrásico, y alejado del discurso cotidiano. A veces hace su aparición en el lenguaje ordinario, como por ejemplo cuando decimos: "Si tú eres agente de la CIA yo soy el Papa", que es claramente un caso de condicional material, trivialmente verdadero por falsedad del antecedente. Pero no es muy frecuente utilizar este tipo de enunciados.

Si nosotros adoptamos aquí este uso material del condicional como uso principal es por una razón muy sencilla: es este tipo de condicional, con diferencia, el que más vamos a tener que utilizar en esta asignatura. Pero hay que señalar que conservamos la posibilidad de utilizar otras clases de condicionales, por ejemplo causal, o de relevancia, siempre que lo advirtamos de forma clara, con expresiones como *"esto es causa de aquello"*, *"relevante para aquello"* u otras así. Es decir: podemos suspender ocasionalmente la aplicación de la convención, haciéndolo notar, de forma enteramente análoga a como ocurre con el resto de convenciones que estamos estableciendo aquí.

§ 1.15. Práctica (Uso del condicional material). Proponer 3 ejemplos nuevos de condicionales diversos, indicando el valor de verdad de cada uno de acuerdo con la convención precedente.

§ 1.16. Observación (Uso del bicondicional). Un enunciado *bicondicional* expresa una condición doble, es decir: *dos* condicionales, uno en un sentido y otro en sentido contrario:

> "Una actriz es premiada con el Oscar *si y sólo si* recibe los votos de la mayoría de miembros de la Academia de las Artes y las Ciencias Cinematográficas de Hollywood".

Ello implica dos cosas: (a) que *si* la mayoría de miembros de la Academia vota a una actriz determinada, *entonces* dicha actriz será premiada con el Oscar (independientemente de que después vaya o no a recogerlo). Y (b) que una actriz *sólo* puede recibir el Oscar por este procedimiento, es decir, que una actriz *sólo* puede recibir el Oscar si ha sido votada por la mayoría de miembros de la Academia de las Artes y las Ciencias Cinematográficas de Hollywood. O lo que es lo mismo: que *si* una actriz recibe el Oscar, *entonces* es que ha sido votada mayoritariamente por los miembros de la susodicha Academia.

Hay que tener mucho cuidado en diferenciar los bicondicionales de los condicionales simples. No es lo mismo decir:

"Serás millonario si te toca el premio gordo de la Lotería Primitiva" (2) que decir:

"Serás millonario *si y sólo si* te toca el premio gordo de la Lotería Primitiva" (3)

El enunciado (3) es un bicondicional, que engloba dos condiciones: (a) que si te toca el gordo de la Primitiva, entonces serás millonario. Esta condición es la misma que (2), y es obviamente verdadera. Y una segunda: (b) que serás millonario *sólo si* te toca el gordo de la Primitiva, o dicho de otro modo: que *si* eres millonario *entonces* es que te ha tocado el gordo de la Primitiva. Ésta es completamente distinta de la anterior, y además en este caso es falsa, ya que el gordo de la Primitiva no es el único medio para llegar a ser millonario. Hay otros medios, como por ejemplo el gordo de la Bonoloto, las quinielas, una herencia, o simplemente haciendo negocios.

En el interior del presente Manual nos vamos a encontrar con muchos bicondicionales, y deben ser correctamente identificados y distinguidos de los condicionales sencillos. Generalmente, los enunciados bicondicionales vienen expresados en la forma:

"p si y sólo si q"

donde p y q son dos enunciados cualesquiera. Pero también se pueden formular diciendo:

"p cuando q y sólo cuando q"
"p exactamente en los mismos casos que q"
"p exactamente cuando q"
"p es condición necesaria y suficiente para q"
Etc.

Ni qué decir tiene que sobre los enunciados bicondicionales cae, de rebote, la convención que hemos hecho antes sobre los condicionales simples, por la sencilla razón de que los bicondicionales no son sino una pareja de condicionales conjugados. Así, para que un enunciado bicondicional sea verdadero tienen que serlo a la vez *los dos* condicionales en los que se desglosa. Y para que cada uno de estos a su vez sean verdaderos tiene que ocurrir que *no* puedan tener antecedente verdadero y consecuente falso.

Cuando haya que demostrar la verdad de un bicondicional, de la forma "p si y sólo si q", generalmente empezaremos demostrando *si p entonces q*, y a continuación demostraremos *si q entonces p*, completando con ello la prueba. El orden en que lo hagamos es lo de menos: lo importante es demostrar que son ciertos los dos condicionales involucrados.

En resumidas cuentas, lo que se suele hacer para demostrar la verdad de un enunciado de estas características es: primero, suponer p para demostrar q, estableciendo con ello que *si p entonces q*. A continuación, suponer q para demostrar p, estableciendo de este modo que *si q entonces p*. Y si conseguimos hacer las dos cosas correctamente, la prueba o demostración habrá terminado. A veces existe un atajo más rápido para acortar la prueba, pero mientras no lo encontremos tenemos que seguir los pasos indicados, que es lo más frecuente.

§ 1.17. Práctica (Uso del bicondicional). Transformar en bicondicional uno de los ejemplos propuestos en § 1.15, indicando el valor de verdad del enunciado resultante.

§ 1.18. Convención (Casos críticos). En ocasiones, para razonar acerca de una clase de objetos, nos paramos a pensar en un individuo ideal, en abstracto. Le ponemos un nombre ficticio, y empezamos a discurrir sobre él, sin presuponer ninguna característica particular específica, excepto la pertenencia a la clase en cuestión. Es un *individuo cualquiera* de dicha clase, lo que técnicamente se denomina a veces un *"caso crítico"*.

También se pueden tomar como casos críticos dos o más individuos. Por ejemplo:

> "Supongamos que un hombre viudo, digamos Manuel, con un único hijo, Manolito, se casa con una chica joven, Juanita, huérfana de padre. Supongamos a continuación, que la madre de Juanita, la viuda Juana, se casa a su vez con el hijo de Manuel, es decir con Manolito. Ocurre entonces que Manolito se convierte en suegro de su padre, padrastro de su padre, padrastro de la mujer de su padre, Juana, y al mismo tiempo su hijastro. Juanita pasa a ser suegra de su madre, así como madrastra e hijastra de Manolito. Y Juana es nuera política de su hija y de su yerno, madrastra de este último . . . "

El uso de casos críticos será muy habitual para nosotros en la presente asignatura: "sea $\underline{\alpha}$ una fórmula cualquiera", "sea $\underline{t}$ un término arbitrario", "sea $\mathscr{L}$ un lenguaje de primer orden", . . . Tendremos que prestar atención para tomar cada caso crítico como un caso *cualquiera*, sin presuponer ninguna característica individual que no sean las expresamente indicadas según se va describiendo el caso.

Además de esto, vamos a adoptar una convención especial: cuando se introducen como casos críticos dos o más objetos pertenecientes a cierta clase, no debemos presuponer que se trata de objetos distintos, a menos que así se indique expresamente. Por ejemplo, cuando digamos "sean $\underline{\alpha}$ y $\underline{\beta}$ fórmulas cualesquiera, o "sean $\underline{\alpha}$ y $\underline{\beta}$ fórmulas", se entenderá que pueden ser la misma. Y sólo cuando se especifique, por ejemplo, "sean $\underline{\alpha}$ y $\underline{\beta}$ fórmulas *distintas*", se utilizará el hecho de que se tiene que tratar de dos fórmulas diferentes.

§ 1.19. Convención (Uso de la cuantificación existencial). Una cuantificación existencial es un enunciado en el que se afirma la existencia de un objeto que satisface cierta condición. La expresión canónica de este tipo de enunciados utiliza el verbo "existir", pero también se pueden formular de otras maneras:

"Existen montañas que superan los 8.000 metros de altura"
"Hay montañas que superan los 8.000 metros de altura"
"Se dan montañas que superan los 8.000 metros de altura"
"Algunas montañas superan los 8.000 metros de altura"
Etc.

La convención que nosotros vamos a adoptar sobre este tipo de enunciados es la siguiente: consideraremos que un enunciado existencial es verdadero cuando exista al menos un objeto que satisfaga la condición en cuestión, y falso en el único caso en que no exista ninguno. A estos efectos, por lo tanto, trataremos de la misma manera a los existenciales que van formulados en singular y a los que van en plural: será suficiente, bajo nuestra convención, con que haya *un* solo objeto que satisfaga la condición exigida.

Es por ello que:

"Existe una montaña que supera los 8.800 metros de altura "
"Hay una montaña que supera los 8.800 metros de altura"
"Alguna montaña supera los 8.800 m. de altura"

y

"Existen montañas que superan los 8.800 metros de altura"
"Hay montañas que superan los 8.800 metros de altura"
"Algunas montañas superan los 8.800 metros de altura"

son todos ellos *verdaderos* por igual, aunque sólo haya 1 montaña, el Everest, que supere dicha altura. Se entiende, naturalmente, que estamos hablando de montañas en la superficie de la Tierra, y de su altura medida sobre el nivel del mar.

Por otra parte, también un caso como

"Hay un chino en China"

será verdadero, de acuerdo con nuestra convención, con independencia de que en China no haya sólo un ciudadano chino, sino 1.300 millones.

Una excepción importante: sólo escapan a esta convención aquellos enunciados existenciales numéricos en los que se especifique claramente que la condición en cuestión es cumplida por un número exacto de objetos distintos, o por una cantidad determinada de objetos dentro de ciertas cotas. En esos casos el existencial será verdadero o falso según que haya efectivamente objetos que cumplan la condición, en el número especificado, o dentro de las cotas especificadas.

Así por ejemplo, los enunciados

"Hay *al menos dos* montañas distintas que superan los 8.000 metros de altura"
"Hay *más de dos* montañas que superan los 8.000 metros de altura"

son ambos verdaderos, porque hay concretamente 14 picos diferentes que superan esa altura. Es decir: es cierto que hay al menos 2, y es cierto que hay más de 2.

Pero

"Hay *al menos dos* montañas distintas que superan los 8.800 metros de altura"

es *falso*, porque de hecho no hay dos, sino sólo una (el Everest, como ya hemos dicho: el pico que le sigue en altura, el llamado "K2", tiene 8.600).

Y por su parte, los enunciados

"Hay *exactamente dos* montañas que superan los 8.000 metros de altura"
"Hay *exactamente una* montaña que supera los 8.000 metros de altura"

son también falsos, porque el número exacto de montañas que superan los 8.000 metros no es ni 2 ni 1, sino 14.

§ 1.20. Práctica (Uso de la cuantificación existencial). Poner dos nuevos ejemplos de existenciales no numéricos, uno verdadero y otro falso. Repetir el ejercicio con existenciales numéricos.

§ 1.21. Convención (Uso de la cuantificación universal). Finalmente, una cuantificación universal es un enunciado en el que se afirma que todos los objetos de cierta clase satisfacen una determinada condición:

"Todos los hombres son mortales"
"Cualquier hombre es mortal"

o sencillamente

"Los hombres son mortales"

que viene a ser sinónima de las anteriores.

Es evidente que un universal es verdadero cuando, efectivamente, *todos* los objetos de la clase en cuestión cumplen la condición correspondiente, y falso en caso contrario. Pero ¿qué ocurre si *no hay* objetos de la susodicha clase? Nuestra convención será, a este respecto, considerar verdaderos a todos aquellos enunciados universales que se refieran a una clase vacía de objetos. Es lo que técnicamente denominamos *"universal vacuamente verdadero"*.

Así por ejemplo,

> "Todas las óperas que compuso Cervantes se han representado en el Gran Teatro del Liceo de Barcelona"

es vacuamente verdadero, y por lo tanto *verdadero* a todos los efectos, por la sencilla razón de que Cervantes no compuso ópera alguna, entendiendo "ópera" en su acepción moderna, naturalmente, esto es, como *drama musical*. Y lo mismo en casos similares.

Por otra parte, y esto también es importante, cuando un enunciado universal se refiera a diversos objetos de una misma clase, no debemos limitarnos a considerar relaciones entre objetos distintos, a menos que el enunciado así lo indique expresamente. Por ejemplo, el universal

> "Con cualesquiera baloncestistas a, b, c, d y e, se puede formar un equipo"

es *falso*, porque si resulta que, por ejemplo, a y b son el mismo, entonces ya no tenemos a 5 jugadores, sino a 4, y es por tanto insuficiente para constituir un equipo de baloncesto conforme al reglamento. Para que fuera verdadero, el universal, debería decir:

> "Con cualesquiera baloncestistas *distintos* a, b, c, d y e, se puede formar un equipo".

§ 1.22. Práctica (Uso de la cuantificación universal). Poner un nuevo ejemplo de universal vacuamente verdadero. Poner un nuevo ejemplo de universal con referencia a grupos de objetos de una misma clase, especificando que se trata de objetos distintos. Determinar los valores de verdad de los dos ejemplos propuestos, de acuerdo con la convención precedente.

§ 1.23. Observación (Combinación de las convenciones). En el enunciado

> "Si hay una estrella en el firmamento, entonces todas las verdes ideas duermen furiosamente"

se hayan involucradas varias de las convenciones establecidas aquí. Un examen detenido muestra que, aplicándole todas las convenciones que se combinan en él, se trata de un enunciado verdadero.

§ 1.24. Práctica (Combinación de las convenciones). Idear otro ejemplo de enunciado complejo, en el que se mezclen varias de nuestras convenciones. Indicar cuál es su valor de verdad, razonando la respuesta paso por paso.

Nociones de teoría de conjuntos

§ 1.25. Definición (Conjunto, pertenencia). Un *conjunto* es una colección de objetos que, considerada como un todo, forma a su vez un nuevo objeto. El ejemplo más típico es el del *rebaño de ovejas*. Un rebaño de ovejas constituye, como tal rebaño, un nuevo objeto, distinto de cada una de las ovejas que lo componen: un objeto que consiste, precisamente, en la colectividad que forman en simultáneo todas ellas.

Dicho esto, la *pertenencia* no es más que la conexión que liga a un objeto con aquellos conjuntos de objetos de los que forma parte. Así, si a es un objeto cualquiera y C un conjunto al que a pertenece, entonces decimos que a es *"miembro"* o *"elemento"* de C, lo cual escribimos de manera abreviada poniendo

$$\text{“}\, a \in C \,\text{”}$$

Y por su parte, para indicar lo contrario, esto es, para indicar que a *no es* un elemento de C, escribimos simplemente: "$a \notin C$".

Por lo demás, para describir conjuntos utilizamos la conocida notación de llaves, como por ejemplo en

$$\text{“}\, C = \{\, a, b \,\} \,\text{”}$$

que significa que C es un conjunto cuyos únicos elementos son a y b. O en:

$$\text{“}\, E = \{\, \text{ciudadanos españoles} \,\} \,\text{”}$$

que significa que E es el conjunto formado por todos los ciudadanos españoles.

Y finalmente, si a, b y c, por ejemplo, son elementos de un conjunto D, entonces utilizaremos también la notación abreviada

$$\text{“}\, a, b, c \in D \,\text{”}$$

para indicar que, en efecto, todos ellos pertenecen a D.

§ 1.26. Definición (Conjuntos disjuntos). A veces dos conjuntos comparten algunos de sus elementos, es decir, tienen elementos comunes. En otros casos no. Cuando dos conjuntos no tienen ningún elemento en común, se dice que son *"conjuntos disjuntos"*.

§ 1.27. Definición (Inclusión). Por otra parte, decimos que un conjunto C *"está incluido"* en otro D, cuando *todos* los elementos de C son elementos de D. En este caso decimos también que C *"es un subconjunto de D"*, y lo escribimos abreviadamente poniendo:

$$\text{“}\, C \subseteq D \,\text{”}$$

Para indicar lo contrario, es decir, para indicar que C *no está* incluido en D, ponemos "$C \nsubseteq D$". Esto puede ocurrir porque sólo sean elementos de D *algunos* de los elementos de C, pero no todos, o bien porque C y D no tengan ningún elemento en común, es decir, porque C y D sean conjuntos disjuntos.

Por ejemplo, si consideramos los siguientes conjuntos:

$$\begin{array}{lcl} B & = & \{\text{ciudadanos de Bélgica}\} \\ H & = & \{\text{ciudadanos de Holanda}\} \\ L & = & \{\text{ciudadanos de Luxemburgo}\} \\ N & = & \{\text{ciudadanos del Benelux}\} \end{array}$$

resulta obvio —dado que el Benelux es un consorcio entre los tres anteriores países— que:

$$B \subseteq N \qquad H \subseteq N \qquad L \subseteq N$$

Mientras que, por ejemplo,

$$B \nsubseteq H \qquad H \nsubseteq L \qquad L \nsubseteq B$$

ya que *no todos* los ciudadanos belgas son holandeses, ni todos los holandeses luxemburgueses, ni estos últimos belgas, aunque pueda haber algunos ciudadanos en particular que tengan varias de estas nacionalidades.

Obsérvese además que, trivialmente, todo conjunto de objetos está incluido en sí mismo, es decir, es un subconjunto de sí mismo. Con lo que en particular, en este caso:

$$B \subseteq B \qquad H \subseteq H \qquad L \subseteq L \qquad N \subseteq N$$

Precisamente a los subconjuntos de un conjunto dado que *no* coinciden con él mismo, se les distingue de una manera especial llamándoles *"subconjuntos propios"*. Por consiguiente, B, H y L son ejemplos de subconjuntos propios de N, mientras que el mismo N no lo es.

§ 1.28. Observación (Inclusión y pertenencia). Teniendo en cuenta que un conjunto es una colección de objetos que constituye a su vez un nuevo objeto, como tal objeto puede pertenecer a otros conjuntos, y ello complica ligeramente las cosas.

Por ejemplo, tomemos los tres conjuntos B, H y L que hemos definido en la sección anterior, y formemos con ellos un nuevo conjunto J, que los tenga a los tres como elementos:

$$J = \{B, H, L\}$$

¿Son iguales el conjunto N antes definido y este J? ¡No, en absoluto! No hay más que fijarse en que el conjunto N tiene millones de elementos, tantos como ciudadanos del Benelux, mientras que el conjunto J tiene sólo 3. Además, los elementos del conjunto N son personas, mientras que los elementos del conjunto J son otros conjuntos, lo cual es muy distinto.

Lo que sí es cierto que los elementos de N son *elementos de conjuntos que son elementos de J*, es decir, son *elementos de los elementos* de J. Pero esto ya es una relación indirecta, que no nos tiene que confundir sobre la respectiva identidad de los dos conjuntos involucrados. Tampoco los hijos de mis hijos son mis hijos, sino mis nietos, valga la comparación.

Recordemos una vez más que cada conjunto consiste en la formación de un nuevo objeto a partir de una determinada colección de objetos, que son sus elementos. Por consiguiente, si la colección de objetos considerada es distinta, esto es, si los elementos son distintos, entonces los conjuntos formados tienen que ser también forzosamente distintos.

Por todo ello podemos decir que B *pertenece a* J ($B \in J$), porque es uno de sus elementos, en efecto, pero *no* que B *pertenezca* a N, porque los elementos de N son personas, y no conjuntos. Es decir: tenemos que $B \notin N$.

Y por su parte, B *está incluido en* N pero *no* en J, ya que *no es cierto* que todos los elementos de B, esos millones de ciudadanos belgas, sean elementos de J, que sólo tiene 3. En definitiva, $B \subseteq N$, pero $B \nsubseteq J$.

§ 1.29. Práctica (Inclusión y pertenencia). Poner un ejemplo similar al de la Observación precedente, en el que varios conjuntos estén *incluidos* en un tercero, y a su vez sean *elementos* de otro distinto. Examinar detalladamente las relaciones de inclusión y pertenencia que se dan entre unos y otros.

§ 1.30. Definición (Unión). Si C y D son conjuntos cualesquiera, entonces la *unión* de C y D, representada por

$$\text{“}C \cup D\text{”}$$

es aquel conjunto cuyos elementos son exactamente todos los de C más todos los de D.

Por ejemplo, si A es el conjunto de todos los habitantes de la provincia de Cáceres, y B es ahora el conjunto de todos los habitantes de Badajoz, entonces claramente

$$A \cup B$$

es el conjunto de todos los habitantes de Extremadura, dado que esta región se compone precisamente de las dos antedichas provincias.

§ 1.31. Observación (Axioma de extensionalidad). En el año 1902 el gran lógico y filósofo inglés Bertrand Russell demostró que la definición incontrolada de conjuntos podía dar lugar a contradicciones. Lo hizo mediante la derivación de la famosa paradoja que lleva su nombre, y a partir de ese momento se hizo patente la necesidad de someter la formación de conjuntos a restricciones especiales.

Nosotros no vamos a entrar aquí en la discusión de esas restricciones, salvo en un único caso, el *axioma de extensionalidad*, que simplemente dice lo siguiente: *“dos conjuntos que tengan exactamente los mismos elementos son iguales, es decir, son uno y el mismo conjunto”*.

Por ejemplo, el conjunto de animales con dos pies y carentes de plumas, y el conjunto de los hombres, es el mismo conjunto, según este axioma, porque de hecho los únicos animales bípedos e implumes que existen somos los hombres. En otras palabras: el conjunto de *animales bípedos e implumes* resulta tener de hecho los mismos elementos que el conjunto de *hombres*, lo cual implica, según el axioma de extensionalidad, que estos dos conjuntos son *iguales*, es decir: son *uno solo.*

Ello supone una restricción indirecta a la formación de conjuntos, debido a que impide postular la existencia de dos conjuntos distintos, atendiendo a cualidades distintas, si resulta que de hecho los elementos de ambos coinciden. Cuando se produzca una tal coincidencia estaremos obligados, de acuerdo con el axioma de extensionalidad, a considerar que los dos conjuntos introducidos son en realidad uno solo.

§ 1.32. Definición (Conjuntos unitarios). Un *conjunto unitario* es aquel que sólo tiene *un* elemento. Por ejemplo, el conjunto

$$C = \{\, \text{ciudades españolas con más de 3 millones de habitantes} \,\}$$

tiene un solo elemento: Madrid, que es la única ciudad española cuya población excede de 3 millones.

Es por ello que podemos describir C más sencillamente poniendo

$$\text{“} C = \{\, \text{Madrid} \,\} \text{”}$$

ya que Madrid es su único elemento. C es obviamente un conjunto unitario.

No hay que confundir a C con la ciudad de Madrid, son cosas muy distintas: Madrid es una ciudad, una entidad compleja, que incluye un territorio, edificios, una infraestructura administrativa y de gobierno, etc. Mientras que C es sencillamente un conjunto con un único elemento.

Tampoco hay que confundir a C con el *conjunto de habitantes de la ciudad de Madrid*: este último tendrá más de 3 millones de elementos, mientras que C sólo tiene 1, así que la diferencia entre ambos es evidente.

En general, siempre que tratemos con un conjunto unitario debemos prestar atención para no confundirlo con el único elemento que contiene.

§ 1.33. Definición (Conjunto vacío). Por su parte, el *conjunto vacío* es aquel que no tiene ningún elemento. Es decir: como su propio nombre indica, es un conjunto *vacío de elementos*.

Puede parecer raro hablar de conjuntos que sólo tienen 1 elemento, y más aún de un conjunto que no tiene ninguno. La idea inicial de conjunto como *colección de cosas* resulta un tanto forzada, en cuanto que ahora estamos considerando *colecciones* de una sola cosa, e incluso de ninguna cosa. En este sentido tanto la presente definición como la anterior se pueden tomar como complementos adicionales a la *Definición* § 1.25 (p. 30), en la que se establecían las bases de las nociones de *conjunto* y *pertenencia*.

Si pensamos en los conjuntos como cajas de cartón conteniendo objetos, por poner una imagen metafórica, entonces un conjunto unitario sería aquella caja que sólo contiene 1 objeto, y el conjunto vacío sería una caja que no contiene objeto alguno.

Y venimos hablando de “el” conjunto vacío porque resulta que solamente hay uno. En efecto, siendo A y B conjuntos vacíos, sin ningún elemento, resulta *trivialmente verdadero* que *todos* los elementos de A son elementos de B y viceversa, ya que no hay ningún elemento ni en A ni en B. O dicho de otro modo: resulta trivialmente verdadero que A y B tienen *exactamente los mismos elementos*. Aplicando el axioma de extensionalidad, por consiguiente, tenemos inmediatamente que A y B son iguales, es decir, son uno y el mismo conjunto.

Ello ocurre con “todos” los conjuntos vacíos, es evidente, por lo que concluimos que son todos iguales entre sí. En otras palabras: que nada más que hay uno. Y para designar este conjunto tan singular le hemos atribuido un nombre propio, “$\varnothing$”.

En definitiva —y atención a esto— de ahora en adelante, siempre que hablemos de un conjunto cualquiera, sea de fórmulas, de términos o de lo que sea, habremos de considerar, a menos que se especifique lo contrario, la posibilidad de que se trate del conjunto vacío $\varnothing$, es decir: del *conjunto vacío* de fórmulas, del *conjunto vacío* de términos, etc.

§ 1.34. Observación (Problema 1). Este es el momento apropiado para plantearse la respuesta al *Problema 1* del listado de *Problemas propuestos* (p. 12).

§ 1.35. Convención (El conjunto de los números naturales). Otro conjunto con nombre propio es el conjunto de los números naturales

$$\mathbb{N} = \{ 0, 1, 2, 3, \dots \}$$

conjunto que va a tener un protagonismo nada desdeñable en la presente asignatura, y al que siempre denotaremos a partir de ahora con la letra indicada.

En teoría de conjuntos se describe la forma de identificar a su vez cada uno de los números naturales con un conjunto específico, representante suyo, que se comporta exactamente igual que él y conserva todas sus propiedades. Ello significa que la estructura de los números naturales se puede contemplar como inmersa en el universo más general de los conjuntos, y por lo tanto, que gran parte de lo que digamos aquí sobre $\mathbb{N}$ afectará también a dicho universo, en el cual aparece representado $\mathbb{N}$ como un pequeño fragmento suyo.

Nosotros, sin embargo, no vamos a entrar aquí a estudiar esa forma de representación. Por lo tanto, trataremos a los números naturales de forma intuitiva y directa, tal y como aprendimos a conocerlos en la Educación Primaria.

§ 1.36. Convención (Expresiones numerarias). También a partir de ahora vamos a empezar a utilizar, en distintos contextos y con significados no siempre idénticos, ciertas expresiones que podríamos llamar "expresiones numerarias". Estas expresiones son *"1-ario"*, *"2-ario"*, *"3-ario"*, *"4-ario"*, *"5-ario"*, *"6-ario"*, y así sucesivamente. En general, por lo tanto, diremos *"n-ario"* para cualquier número natural $n \geqslant 1$.

Algunas de estas expresiones tienen nombre propio, como por ejemplo:

"monario",	que es sinónimo de	"1-ario"
"binario",	"	"2-ario"
"ternario",	"	"3-ario"
"cuaternario",	"	"4-ario"

Hay más expresiones numerarias con nombre propio, pero estos cuatro primeros son los más habituales, y los únicos que se utilizarán en la presente asignatura.

§ 1.37. Convención (Propiedades y relaciones). Hay una diferencia evidente entre las propiedades y las relaciones, que ya se mencionó de pasada en § 1.9, y que estriba en lo siguiente: una propiedad es una condición que afecta de manera singular a un objeto, mientras que una relación involucra cierta situación respectiva entre dos o más objetos.

Así por ejemplo, *ser ebanista* es una propiedad, porque es una condición que puede ser cumplida o incumplida por cada persona individualmente considerada. Mientras que *ser marido de* involucra necesariamente la consideración de dos personas: por una parte el marido, y por otra la mujer con la que el marido esté casado.

Decimos que la relación *ser marido de* es *"binaria"* porque requiere para su consideración de *dos* objetos. Hay también relaciones *ternarias*, que requieren a 3 objetos. Por ejemplo *estar en medio de*, que involucra, por una parte, al objeto que está en medio, y por otra a los dos objetos que están a cada uno de sus lados. También hay relaciones *cuaternarias*, 5-arias, 6-arias, etc.

Las propiedades, por su parte, se pueden contemplar como casos especiales de relaciones, en las cuales el número de objetos involucrados es exactamente 1. Es decir, se puede considerar que una propiedad es una *relación monaria*. De hecho, esto va a resultar muy conveniente para nosotros en numerosas ocasiones a lo largo de este Manual, por lo que a partir de ahora, siempre que hablemos de "relaciones monarias" se entenderá que se trata de propiedades. Y siempre que hablemos de "relaciones" sin más, se entenderá que incluye a las relaciones monarias, esto es, a las propiedades.

En otras palabras: que adoptamos a partir de aquí una nueva manera de llamar y clasificar a las propiedades, a saber: como "relaciones monarias". Sin perjuicio de que podamos seguir llamándolas "propiedades" cuando nos interese.

§ 1.38. Definición (Secuencias). Una *secuencia* es una lista de objetos en la que estos aparecen en un cierto orden, y con posibles repeticiones. Cada aparición de un objeto en una secuencia constituye un *componente* de la misma.

Un ejemplo inmediato es la misma serie de los números naturales, tomados en su orden habitual:

$$0, 1, 2, 3, \ldots$$

Otro ejemplo es la secuencia que conforma el número de teléfono de mi despacho:

$$9, 6, 8, 3, 6, 7, 7, 5, 3$$

En otras secuencias pueden aparecer objetos diversos, no necesariamente números.

Toda secuencia ha de tener un *primer componente*, que es el objeto con el que se inicia la lista. Si no tiene más que ése, decimos que se trata de una *"secuencia monaria"*. Si tiene 2 componentes, decimos que es una *"secuencia binaria"*. Si tiene 3, *"ternaria"*. Y así sucesivamente. Por ejemplo, la secuencia que conforma el número de teléfono de mi despacho es una secuencia 9-aria. Y en general, para cada número natural n mayor o igual a 1, una *secuencia n-aria* es aquella que tiene exactamente n componentes, es decir, que consta de n lugares en la lista.

Finalmente, a las secuencias que no tienen un último componente, porque no se acaban nunca, como le ocurre a la serie de los números naturales, las llamamos *"secuencias infinitas"*.

§ 1.39. Convención (Relaciones y secuencias). Según lo dicho, es evidente que las propiedades (o relaciones monarias) involucran secuencias monarias de objetos, las relaciones binarias involucran secuencias binarias, y así sucesivamente.

Pues bien: si R es una relación n-aria para algún número natural $n \geqslant 1$, y

$$a_1, a_2, \ldots, a_n$$

es una secuencia n-aria de objetos, entonces pondremos

$$\text{“}R(a_1, a_2, \ldots, a_n)\text{”}$$

para significar que los objetos $a_1, a_2, \ldots, a_n$, *tomados en ese orden*, cumplen la relación R. Y pondremos

$$\text{“}\sim R(a_1, a_2, \ldots, a_n)\text{”}$$

(leído *"no $R(a_1, a_2, \ldots, a_n)$"*), para denotar lo contrario.

§ 1.40. Práctica (Relaciones y secuencias). Poner un ejemplo de relación monaria R, otro de relación binaria S y otro de relación ternaria T . Buscar secuencias en las que se cumplan cada una de ellas, y otras en las que no se cumplan. Describir todos los casos utilizando la notación que se acaba de introducir.

§ 1.41. Definición (Función). Una *función* es una operación que al ser aplicada a uno o más objetos, da como resultado otro objeto. Por ejemplo la función *cuadrado de*, al ser aplicada al número 2, nos da como resultado el número 4. Y la función *dividido por*, al ser aplicada a los números 56 y 7, por ese orden, nos da como resultado el número 8. En otras palabras: 8 es el resultado de efectuar la operación *56 dividido por 7*.

Los objetos a los que se aplica una función se llaman *"argumentos de la función"*, y el resultado que se obtiene como consecuencia se denomina *"valor que toma la función para esos argumentos"*. Así, la función *cuadrado de*, al ser aplicada al número 2 como argumento, nos da el número 4 como valor. Y la función *dividido por*, al ser aplicada a los argumentos 56 y 7, por ese orden, nos da como valor el número 8.

Una función es *monaria* cuando se aplica a un solo argumento, como por ejemplo la función *cuadrado de*. Una función es *binaria* cuando se aplica a secuencias de dos argumentos, como por ejemplo la función *dividido por*. Hay también funciones *ternarias*, *cuaternarias*, y así sucesivamente. En general, para cada natural $n \geqslant 1$, una función n-aria es aquella que se aplica a secuencias n-arias de objetos.

Ni qué decir tiene que el orden en que aparezcan los argumentos en la secuencia suele ser muy importante. No es lo mismo

$$56 \text{ dividido por } 7$$

que

$$7 \text{ dividido por } 56$$

En el primer caso el resultado es 8, mientras que en segundo es 0.125. En este sentido, hay que poner atención en identificar al primer argumento, al segundo argumento, al tercer argumento (si lo hay), etc., en el orden exacto en que conforman la secuencia que se presenta a la función para su aplicación.

No sólo hay funciones numéricas, como las que acabamos de citar, sino también las hay que relacionan objetos de otro tipo. Por ejemplo, la función *signo del Zodiaco de* atribuye a cada persona un signo del Zodiaco, atendiendo a su fecha de nacimiento. Y la función *lugar de empadronamiento de* atribuye a cada persona un municipio, aquél en que se encuentra empadronado.

Lo que sí es importante señalar es que, *para los propósitos del presente Manual*, consideraremos que una función no puede dar nunca dos valores distintos cuando se aplica a un mismo objeto o secuencia de objetos. Es decir: el valor de una función para un argumento, si lo hay, habrá de ser único. La operación *color de*, por ejemplo, que asigna a cada objeto el color de su superficie, no será una función para nosotros, según lo que acabamos de indicar, porque con frecuencia las cosas tienen más de un color diferente. Como tampoco será función para nosotros la operación *hijo de*, que asigna a cada persona sus hijos, porque muchas personas tienen más de uno, etc.

Finalmente, si f es una función n-aria para algún número natural $n \geqslant 1$, y $a_1, a_2, \ldots, a_n$ es una secuencia n-aria de objetos, entonces escribimos

$$\text{“}f(a_1, a_2, \ldots, a_n)\text{”}$$

para denotar el resultado obtenido al aplicar la función f a los objetos de la secuencia $a_1, a_2, \ldots, a_n$, por ese orden.

§ 1.42. Práctica (Funciones). Poner dos nuevos ejemplos de funciones numéricas, una monaria y otra binaria. Aplicarlas a algún argumento o argumentos, determinando el valor que se obtiene en cada caso.

§ 1.43. Definición (Dominio de una función). Lo que sí puede ocurrir al aplicar una función a un argumento o cadena de argumentos, es que no obtengamos valor alguno. En estos casos decimos que la función en cuestión *“no está definida”* para ese argumento o argumentos.

Por ejemplo, cuando la función *empadronado en* se aplica a un niño recién nacido no produce nada, porque el niño no está empadronado aún en ningún sitio. O cuando la función *raíz cuadrada de* se aplica a una silla, tampoco produce nada.

Precisamente llamamos *“dominio”* de una función, a aquel conjunto de objetos, o de secuencias de objetos, para los cuales dicha función sí está definida. Así por ejemplo: la función *lugar de empadronamiento de* tiene como dominio a las personas que estén registradas en alguna Oficina de Empadronamiento, pero no a los gatos. Y la función *cuadrado de* tiene como dominio a los números de los diversos campos numéricos, pero no a los muebles ni a otros objetos.

§ 1.44. Definición (Correspondencia biunívoca, cardinalidad). Es obvio que hay conjuntos más grandes que otros. A la teoría de conjuntos le interesa medir comparativamente los tamaños de los conjuntos, y para ello utiliza una noción llamada *“cardinalidad”*.

Decimos que dos conjuntos tienen *“la misma cardinalidad”* cuando sus elementos se pueden poner en *correspondencia biunívoca*, es decir: cuando sus elementos se pueden emparejar de tal forma que cada elemento de un conjunto tenga una única y exclusiva pareja en el otro, distinta de la que tienen los demás, y sin que quede ningún elemento por emparejar en ninguno de los dos conjuntos.

Establecer una correspondencia biunívoca entre los elementos de dos conjuntos viene a ser algo así, metafóricamente hablando, como “formar parejas de baile” con los elementos de uno y otro, de forma que, una vez efectuadas las parejas, no quede ningún elemento suelto en ninguno de los dos, y pueda empezar el baile.

Supongamos que una peña de amigos, integrada por los caballeros Ambrosio, Agapito, Ruperto, Eulogio y Emeterio, se reúnen en una fiesta con las damas Venancia, Engracia, Balbina, Gertrudis y Lucrecia. Es obvio que en este caso se puede establecer una correspondencia biunívoca entre el conjunto de hombres y el conjunto de mujeres de la fiesta. Por ejemplo:

Ambrosio	———	Venancia
Agapito	———	Engracia
Ruperto	———	Balbina
Eulogio	———	Gertrudis
Emeterio	———	Lucrecia

Hay otras formas de hacerlo, es decir, hay otras formas de establecer una correspondencia biunívoca entre estos dos conjuntos. Pero lo importante es que haya *al menos una*, y ello demuestra que se trata de dos conjuntos con la misma cardinalidad.

Por otra parte, si dados los conjuntos cualesquiera A y B, se puede establecer una correspondencia biunívoca entre A y *un subconjunto propio* de B, *pero no se puede establecer entre A y todo B*, entonces es que A tiene una *cardinalidad menor* que B.

Por ejemplo, si en una segunda reunión de las dos peñas de amigos, resulta que a última hora falta Emeterio,

Ambrosio	————	Venancia
Agapito	————	Engracia
Ruperto	————	Balbina
Eulogio	————	Gertrudis
		Lucrecia

entonces se puede poner en correspondencia biunívoca el conjunto de hombres asistentes a esta segunda fiesta con un subconjunto propio del conjunto de mujeres, pero no con todo él, como es evidente. El conjunto de hombres asistentes a esta segunda reunión, por consiguiente, tiene una cardinalidad menor que la del conjunto de mujeres.

§ 1.45. Definición (Conjuntos enumerables y supernumerables). La razón de medir los tamaños de los conjuntos por el procedimiento de las correspondencias biunívocas, en apariencia tan aparatoso, reside en el tratamiento de los conjuntos infinitos, para los cuales sólo se puede seguir este tipo de estrategias. Fue el matemático alemán de origen ruso, Georg Cantor, creador de la moderna teoría de conjuntos, el primero en elaborar un tratamiento sistemático para su aplicación.

Así por ejemplo, podemos verificar que el conjunto $\mathbb{N}$ de los números naturales tiene la misma cardinalidad que el conjunto de los números pares (!):

0	————	0
1	————	2
2	————	4
3	————	6
4	————	8
5	————	10
	...	

Ésta es, sin lugar a dudas, una correspondencia biunívoca entre ambos conjuntos: se van formando parejas distintas, sin que puedan sobrar elementos en ninguno de los dos. "Puede dar comienzo el baile", por seguir con la metáfora apuntada antes.

En general, a todos los conjuntos que tienen la misma cardinalidad que el conjunto $\mathbb{N}$ de los números naturales, que son muchos, se les llama *"conjuntos enumerables"*. Es decir: un *conjunto enumerable* es aquel que tiene la misma cardinalidad que el conjunto de los números naturales.

Todos los conjuntos enumerables son obviamente conjuntos *infinitos*. Son, en particular, los más pequeños de los conjuntos infinitos: los "pipiolos infinitos", podríamos decir.

Por su parte, a los conjuntos cuyo tamaño es mayor que $\mathbb{N}$, que también los hay, los llamamos *"conjuntos supernumerables"*.

Y por último, los conjuntos *finitos* serán, evidentemente, aquellos de cardinalidad aún menor que $\mathbb{N}$. Es decir, los conjuntos finitos serán aquellos que se pueden poner en correspondencia biunívoca con un subconjunto propio de $\mathbb{N}$, pero *no* con todo $\mathbb{N}$. El conjunto de los días de la semana, por ejemplo, o el conjunto de los ciudadanos chinos, etc.

§ 1.46. Práctica (Conjuntos enumerables). Proponer otros dos ejemplos distintos de conjuntos enumerables. Razonar la respuesta.

§ 1.47. Práctica* (Argumento de la diagonal de Cantor). Sea $\mathbb{N}'$ el conjunto de todas las secuencias infinitas de números naturales. Probar que $\mathbb{N}'$ es un conjunto supernumerable. Para ello, tomar una correspondencia cualquiera, y mostrar que no puede ser biunívoca, modificando sistemáticamente aquella secuencia formada en diagonal a partir del primer componente de la pareja del 0.

La inducción matemática

§ 1.48. Observación (La inducción matemática). La *inducción matemática* es el hecho estructural más interesante que caracteriza a los números naturales. Se puede formular en dos versiones, el llamado *"principio de inducción débil"* y el *"principio de inducción fuerte"*. A pesar de sus nombres, se trata de principios estrictamente equivalentes.

Aunque a primera vista pueda parecer que este principio matemático no tiene ninguna relación con la lógica, el hecho es que su uso resulta de tal utilidad y eficacia, que hoy por hoy es sin duda la principal herramienta que utiliza el lógico para su trabajo, comparable al uso que hacen los químicos de la regla de tres, o al empleo que hace de la llave inglesa un fontanero. A veces aparece presentado en otros formatos, como el llamado "principio del número menor", o bien como una estructura de instrucciones recursivas, pero esencialmente se trata de un mismo principio de razonamiento.

Por lo demás, la inducción matemática debe ser cuidadosamente distinguida de la llamada *"inducción empírica"*, que se estudia en epistemología y en filosofía de la ciencia, y con la que no tiene nada que ver en absoluto. Precisamente para subrayar la diferencia entre ambas se denomina a veces a la inducción matemática, *"inducción completa"*.

§ 1.49. Definición (Función de sucesión). Llamamos *"sucesor"* de un número natural n al número que le sigue en la serie, es decir, a $n+1$. Así, el sucesor de 0 es el 1, el sucesor de 1 es el 2, etc. Todos los números naturales tienen un sucesor, y todos son sucesores de algún número, excepto el 0, naturalmente, que no es sucesor de ninguno.

A la función que asigna a cada natural n su sucesor $n + 1$ la llamamos *"función de sucesión"*, y la representamos con la letra "s". Por lo tanto, para todo $n \in \mathbb{N}$ tenemos

$$s(n) = n + 1$$

§ 1.50. Observación (El principio de inducción débil). Supongamos que el número 0 cumple una determinada propiedad P, y que para cualquier número natural n sucede que

si n tiene la propiedad P, también la tiene su sucesor $s(n)$

En estas circunstancias, parece claro que *todos los números naturales tienen la propiedad en cuestión*. Pues bien: este hecho tan sencillo es justamente lo que se conoce como *"principio de inducción débil"*.

Por ejemplo, si el 0 es un número *espialidoso* —con independencia de lo que signifique esto— y ocurre que el sucesor de cualquier número *espialidoso* es también *espialidoso*, entonces resulta evidente que serán *espialidosos* el 1, el 2, el 3, el 4, ..., y así sucesivamente hasta "barrer" toda la serie de los números naturales.

Se trata, como vemos, de una especie de "efecto dominó", por decirlo de alguna manera: si cae la primera ficha, y cada ficha que caiga hace caer a la siguiente, entonces resulta impepinable que caigan todas.

§ 1.51. Observación (Uso de la inducción débil en lógica). La razón de que el principio de inducción débil tenga tan amplia aplicación en lógica, deriva de la estrecha relación que existe entre los números naturales y las distintas entidades que estudia ésta, de tal manera que los efectos del principio de inducción matemática resultan proyectados sobre aquellas entidades.

Por ejemplo, las fórmulas de un lenguaje formal vienen ordenadas en *grados*, según ciertas características, fácilmente detectables, de cada fórmula: hay fórmulas *de grado* 0 —las llamadas *"fórmulas atómicas"*—, hay fórmulas *de grado* 1, *de grado* 2, y así sucesivamente.

En estas condiciones, supongamos que todas las fórmulas *de grado* 0 tienen una determinada propiedad P. Y que además sucede, para cualquier número natural n, que *si* todas las fórmulas *de grado* n cumplen la propiedad P, *entonces* también la cumplen todas las fórmulas *de grado* $n + 1$. Es claro que en tal caso *todas* las fórmulas de ese lenguaje, sin excepción, cumplirán la propiedad de marras: en efecto, la cumplen todas las fórmulas de grado 0, como la cumplen todas las fórmulas de grado 0, también la cumplirán todas las de grado 1, como la cumplen todas las de grado 1, también la cumplirán las de grado 2, y así hasta el infinito.

§ 1.52. Observación (Pruebas por inducción débil). Como es natural, el principio de inducción débil no afecta a todas las propiedades que quepa plantearse sobre las fórmulas de un lenguaje formal, sino sólo a aquellas que satisfagan las dos condiciones relevantes. Es decir: a aquellas propiedades que sean cumplidas por las fórmulas de grado 0, y tales que, para cualquier número natural n, si son cumplidas por las de grado n, también lo sean por las de grado $n + 1$.

Precisamente una *prueba por inducción débil* consiste en establecer que una determinada propiedad cumple esas dos condiciones, como estrategia demostrativa para concluir, al final, que la propiedad la cumplen *todas* las fórmulas de ese lenguaje formal, sin excepción.

Una prueba por inducción débil tiene, por lo tanto, dos partes. Una primera parte, llamada *"base"*, en la cual se trata de probar que todas las fórmulas de grado 0 cumplen la propiedad en cuestión. Y una segunda parte, llamada *"paso de inducción"*, en la cual se supone que, para un número arbitrario n, todas las fórmulas de grado n tienen dicha propiedad, y a partir de ahí se trata de probar que también las fórmulas de grado $n+1$ la tienen.

A la suposición de que las fórmulas de grado n cumplen la propiedad en cuestión se le llama *"hipótesis de inducción"*. No es una hipótesis definitiva, sino sólo un eslabón en la estructura general de la prueba, que utilizamos para probar que las fórmulas de grado $n+1$ también satisfacen dicha propiedad.

Y en fin, si conseguimos hacer todo esto, la prueba inductiva habrá acabado, habiendo demostrado con ello que la propiedad escogida la cumplen *todas* las fórmulas de ese lenguaje formal.

En el caso de otras muchas entidades de las que trata la lógica no cabe hablar de *"grados"*, pero sí de otras características peculiares distintas, que permiten asignar a cada una un número natural, 0, 1, 2, 3, ..., y someterlas así a operaciones similares a la que acabamos de describir para el caso de las fórmulas.

§ 1.53. Observación (El principio de inducción fuerte). El principio de inducción fuerte tiene una apariencia más sofisticada que el débil, pero en definitiva viene a decir algo muy parecido. Sea P una propiedad numérica, y supongamos que para cualquier número natural n,

si todos los naturales anteriores a n tienen la propiedad P, entonces n también la tiene

En estas condiciones, no es difícil ver que la propiedad P en cuestión será cumplida por *todos* los números naturales. Y este hecho es el que se conoce como *"principio de inducción fuerte"*.

En efecto, supongamos que la propiedad de ser *espialidoso* —con independencia de lo que signifique esto— es una propiedad tal que para cualquier número n,

si todos los naturales anteriores a n son *espialidosos*, también n es *espialidoso*

Lo que ocurre entonces es que el 0 es *espialidoso*, ya que al no haber números naturales anteriores a 0, resulta trivialmente verdadero que "todos los números naturales anteriores a 0 son *espialidosos*".

A continuación, establecido que el 0 es *espialidoso*, podemos afirmar que "todos los naturales anteriores a 1 son *espialidosos*", ya que el único número natural anterior a 1 es el propio 0. Por lo tanto 1 también es *espialidoso*.

Por consiguiente, si 0 y 1 son *espialidosos*, entonces tenemos que todos los naturales anteriores a 2 son espialidosos, con lo que también lo es el 2. Y así sucesivamente. Queda claro pues, que si la propiedad de ser *espialidoso* satisface la condición inicial, entonces todos los números naturales han de ser *espialidosos* sin más remedio.

§ 1.54. Observación (Pruebas por inducción fuerte). El principio de inducción fuerte resulta para la lógica tan útil o más que el primero. La única diferencia importante es que una prueba por inducción fuerte carece de *base*. Consta únicamente de un *paso de inducción*, en el cual, se toma un número natural arbitrario n, se supone, volviendo al ejemplo de las fórmulas, que *todas las fórmulas de grado menor que n tienen la propiedad en cuestión*, y a partir de ahí se trata de demostrar que las propias fórmulas de grado n

también la tienen. La suposición, una vez tomado el número arbitrario n, de que todas las fórmulas de grado menor que n tienen la propiedad escogida, constituye obviamente la *hipótesis de inducción* de las pruebas por inducción fuerte.

En fin, desde el comienzo del próximo Módulo y ya a lo largo de todo el Manual, veremos numerosos ejemplos de cómo la inducción matemática se aplica, en sus dos versiones, a la consecución de demostraciones, así como a la definición de diversas propiedades, relaciones, funciones y conjuntos, que nos ayudarán a caracterizar y a conocer mejor nuestros objetos de estudio.

MÓDULO 2

Elementos de lógica de primer orden

Términos

§ 2.1. Definición (Lenguaje de primer orden). Todo *lenguaje de primer orden* se compone de tres ingredientes distintos: un conjunto de símbolos, que constituye su *alfabeto*, un conjunto de *términos* y un conjunto de *fórmulas.*

La caracterización de cada uno de estos componentes conlleva cierta dificultad, por lo que efectuaremos las definiciones correspondientes por separado.

§ 2.2. Definición (Alfabeto). El *alfabeto* de un lenguaje de primer orden es un conjunto de símbolos cualesquiera, todos ellos distintos entre sí, y clasificados en los siguientes subconjuntos o grupos:

1. Un conjunto de símbolos llamados *"variables"*, que vienen dados en forma de una secuencia infinita:

   $$\underline{v_1},\ \underline{v_2},\ \underline{v_3},\ \ldots$$

   Es decir: una variable $\underline{v_n}$ para cada número natural $n \geqslant 1$.

2. Un conjunto de símbolos llamados *"constantes"*.

3. Para cada número natural $n \geqslant 1$, un conjunto de *símbolos funcionales n-arios.* Es decir:

   - un conjunto de símbolos funcionales monarios
   - un conjunto de símbolos funcionales binarios
   - un conjunto de símbolos funcionales ternarios

   y así sucesivamente.

4. Para cada número natural $n \geqslant 1$, un conjunto de *símbolos relacionales n-arios.* Es decir:

- un conjunto de símbolos relacionales monarios (también llamados *"símbolos predicativos"*)
- un conjunto de símbolos relacionales binarios
- un conjunto de símbolos relacionales ternarios

y así sucesivamente.

5. Un *símbolo de igualdad*, $\underline{=}$
6. Un *símbolo de negación*, $\underline{\neg}$
7. Un *símbolo condicional*, $\underline{\rightarrow}$
8. Un *símbolo de cuantificación universal*, $\underline{\forall}$

Los cuatro últimos símbolos constituyen los *símbolos lógicos* del alfabeto. A los símbolos de negación y condicional los llamamos *"conectivas"*. El símbolo de igualdad se considera también un símbolo relacional binario, pero es especial, al contarse entre los símbolos lógicos.

A las constantes, símbolos funcionales y relacionales distintos del de igualdad, los llamamos *"símbolos peculiares"*. Es decir, exceptuando las variables y los símbolos lógicos, todos los demás son símbolos peculiares.

§ 2.3. Observación (Aspecto de los símbolos). Cada símbolo de un lenguaje de primer orden puede tener cualquier forma gráfica, e incluso sonora u otra, con tal de que sea *reconocible* y lo distinga del resto, es decir: que una vez definido el lenguaje podamos identificar ese símbolo cada vez que aparezca representado en algún sitio. De este modo, cualquier símbolo reconocible puede funcionar como la variable $\underline{v_1}$ de un lenguaje de primer orden, o como la variable $\underline{v_2}$, o como el símbolo de igualdad, etc.

Para entender esto mejor, pensemos en lo siguiente. Cuando decimos, por ejemplo, "el mando directo de una compañía de soldados es el *capitán*" (o: "toda compañía de soldados tiene un mando directo, el *capitán*"), utilizamos la palabra "capitán" para referirnos a aquel oficial al que, por graduación, le corresponde el mando de este tipo de unidad militar. Pero con ello no estamos identificando a ningún capitán en concreto, ni mucho menos describiéndolo físicamente.

Pues bien: del mismo modo, cuando decimos "la primera variable de un lenguaje de primer orden es $\underline{v_1}$" (o: "todo lenguaje de primer orden tiene una primera variable, $\underline{v_1}$"), tan sólo estamos poniendo una etiqueta especial a esa variable, una especie de abreviatura para poder referirnos a ella más cómodamente en el futuro. Pero no estamos especificando qué símbolo es, ni la forma que tiene: en cada lenguaje de primer orden habrá un símbolo desempeñando el papel de variable $\underline{v_1}$, y los símbolos que desempeñen esa función en lenguajes distintos no tienen por qué parecerse en absoluto.

Por poner un ejemplo, la letra eñe (decimoséptima letra del abecedario español), podría servir perfectamente como variable $\underline{v_1}$ de un lenguaje de primer orden, con la única condición de que fuese distinta del resto de símbolos de su alfabeto.

§ 2.4. Observación (El símbolo de igualdad). Según lo dicho, también el símbolo de igualdad de un lenguaje de primer orden puede tener cualquier forma reconocible: no tiene por qué parecerse a la doble barra horizontal, ni a la doble barra horizontal subrayada. La misma letra eñe podría servir como símbolo de igualdad de un lenguaje de primer orden, siempre que dicha letra no desempeñara ya en ese lenguaje otra función distinta.

Ahora bien: lo que sí tiene forma de doble barra horizontal subrayada es *el nombre* abreviado que nosotros hemos dado al símbolo de igualdad de cualquier lenguaje de primer orden. Y ese nombre abreviado *se parece* a su vez a la abreviatura habitual para acortar nuestra expresión castellana "es igual a": en realidad es exactamente igual que esa abreviatura, pero añadiéndole el subrayado.

Esto significa que debemos hacer un esfuerzo especial para no confundirnos, verificando cuidadosamente, cada vez que aparezca una de estas abreviaturas, de cuál de las dos se trata: si se trata de mencionar al símbolo de igualdad de un lenguaje de primer orden, aparecerá la doble barra horizontal subrayada. Y si se trata de acortar la expresión "es igual a", aparecerá la doble barra horizontal sin más. Por ejemplo, podemos abreviar la frase "dos más tres es igual a cinco" poniendo

$$“\, 2 + 3 = 5 \,”$$

pero *no* poniendo

$$“\, 2 + 3 \mathrel{\underline{=}} 5 \,”$$

lo cual significa *dos más tres—símbolo de igualdad de un lenguaje de primer orden—cinco*, y no tiene ningún sentido.

§ 2.5. Observación (Distintos lenguajes de primer orden). Como ya hemos visto, todo lenguaje de primer orden ha de poseer un conjunto enumerable de variables, así como cuatro símbolos lógicos.

En cuanto a los símbolos peculiares, sin embargo, aunque todo lenguaje de primer orden posee, por ejemplo, un conjunto de constantes, ese conjunto puede ser más grande o más pequeño, o incluso ser el conjunto vacío, en cuyo caso el lenguaje en cuestión no poseerá ninguna constante. Y lo mismo ocurre con todos los símbolos funcionales y relacionales, exceptuando el símbolo de igualdad, que siempre ha de estar presente.

En definitiva, de acuerdo con nuestra definición, lo único en que los lenguajes de primer orden pueden diferir unos de otros en cuanto a la estructura de su alfabeto, es en la cantidad y distribución de sus símbolos peculiares.

§ 2.6. Definición (Cardinalidad de un lenguaje). La *cardinalidad de un lenguaje de primer orden* es sencillamente la cardinalidad que tenga el conjunto de símbolos que componen su alfabeto.

Como todo lenguaje de primer orden cuenta con infinitas variables, la cardinalidad mínima de su alfabeto será igual a la de $\mathbb{N}$. Por lo tanto, los lenguajes de primer orden de tamaño menor, serán lenguajes enumerables. Sin embargo, si entre los símbolos peculiares hay, por ejemplo, un conjunto supernumerable de constantes, entonces la cardinalidad del lenguaje será también supernumerable.

§ 2.7. Definición (Cadenas, subcadenas). Una *cadena* de un lenguaje de primer orden $\mathscr{L}$ es una secuencia finita de símbolos, todos los cuales pertenecen al alfabeto de $\mathscr{L}$. Dentro de una cadena, las *ocurrencias* de cada símbolo de $\mathscr{L}$ son todas y cada una de las apariciones de ese símbolo en dicha cadena. Y llamamos *"longitud"* de una cadena al número total de ocurrencias de símbolos que aparecen en ella.

Además, si $\underline{t}$ es una cadena de un lenguaje de primer orden, entonces cualquier fragmento $\underline{r}$ de la cadena $\underline{t}$ es una *subcadena* de $\underline{t}$. Esto incluye la posibilidad de que la *subcadena* $\underline{r}$ sea la misma $\underline{t}$. En concreto, decimos que $\underline{r}$ es una *"subcadena propia"* de $\underline{t}$ cuando sucede que $\underline{r}$ es un fragmento *propio* de $\underline{t}$, es decir, un fragmento de $\underline{t}$ pero no $\underline{t}$ entero.

Finalmente, $\underline{r}$ es una *subcadena inicial* de $\underline{t}$ si consiste en un fragmento inicial de $\underline{t}$, es decir: si el primer símbolo de $\underline{r}$ es el primer símbolo de $\underline{t}$, el segundo símbolo de $\underline{r}$ es el segundo símbolo de $\underline{t}$, y así hasta que termine $\underline{r}$.

Por ejemplo, dado un lenguaje de primer orden que contenga una constante $\underline{d}$, entonces podemos formar la cadena

$$\underline{\forall v_1 \forall\forall d v_1 v_1 d v_1 v_1}$$

que contiene 5 ocurrencias de la variable $\underline{v_1}$, 2 ocurrencias de la constante $\underline{d}$ y 3 ocurrencias del símbolo de cuantificación universal. La longitud de esta cadena es 10, y son subcadenas suyas, entre otras:

$$\underline{\forall v_1 \forall\forall d v_1} \qquad \underline{v_1 d v_1} \qquad \underline{v_1} \qquad \underline{\forall v_1 \forall\forall d v_1 v_1 d v_1 v_1}$$

de las cuales son subcadenas propias las tres primeras, y son subcadenas iniciales la primera y la cuarta.

§ 2.8. Definición (Término). Decimos que una cadena $\underline{t}$ de un lenguaje de primer orden $\mathscr{L}$ es un *"término"* cuando cumple con alguna de estas tres condiciones:

1. $\underline{t}$ es una variable de $\mathscr{L}$.
2. $\underline{t}$ es una constante de $\mathscr{L}$.
3. $\underline{t}$ está compuesta por un símbolo funcional n-ario $\underline{f}$ de $\mathscr{L}$ seguido de otras cadenas $\underline{t_1}, \underline{t_2}, \ldots, \underline{t_n}$, todas las cuales son *términos* de $\mathscr{L}$:

$$\underline{f t_1 t_2 \ldots t_n}$$

   En este caso decimos que los términos $\underline{t_1}, \underline{t_2}, \ldots, \underline{t_n}$ son los *"argumentos"* del símbolo funcional $\underline{f}$.

La definición de "término" es nuestro primer ejemplo de definición por inducción matemática, en este caso por inducción fuerte sobre la longitud de la cadena en cuestión: la definición procede tomando una cadena arbitraria del lenguaje $\mathscr{L}$, la cadena $\underline{t}$, y especificando las condiciones bajo las cuales $\underline{t}$ es un término, dada la suposición de que ya sabemos determinar si son términos o no las cadenas de longitud menor que $\underline{t}$. En particular, nótese que las cadenas $\underline{t_1}, \underline{t_2}, \ldots, \underline{t_n}$ de la cláusula (3), han de tener todas ellas, como es obvio, una longitud menor que $\underline{t}$, ya que $\underline{t}$ contiene una primera ocurrencia del símbolo funcional $\underline{f}$ que no puede estar en ninguna de ellas.

Por ejemplo, si $\mathscr{K}$ es un lenguaje de primer orden que contiene:

- una constante $\underline{d}$
- un símbolo funcional monario $\underline{g}$
- un símbolo funcional binario $\underline{h}$
- un símbolo predicativo $\underline{P}$

entonces las siguientes cadenas son casos claros de términos de $\mathscr{K}$:

$$\begin{array}{rrr} \underline{v_1} & \underline{v_7} & \underline{d} \\ \underline{gv_1} & \underline{gv_7} & \underline{gd} \\ \underline{hv_1v_1} & \underline{hv_1v_7} & \underline{hdd} \\ \underline{hv_1gv_7} & \underline{hv_7gv_7} & \underline{hv_7hdd} \end{array}$$

Como podemos apreciar, los tres primeros ejemplos son del tipo más sencillo posible: una variable sola (como $\underline{v_1}$ o $\underline{v_7}$), o una constante sola (como $\underline{d}$, en este caso). Este tipo de términos resultan muy fáciles de formar y de identificar.

Los términos que presentan algo más de complicación son los que están formados por un símbolo funcional seguido por otros términos. Aquí lo que hay que tener presente es que el símbolo funcional en cuestión sea monario, binario, ternario, etc., para saber por cuántos términos ha de ir seguido. Por ejemplo, un símbolo funcional monario como es el símbolo $\underline{g}$ del lenguaje $\mathscr{K}$, tiene que ir seguido de un solo término, para formar otro término. Y en efecto, así se forman los términos $\underline{gv_1}$, $\underline{gv_7}$ y $\underline{gd}$, que aparecen en la 2ª línea de nuestra lista de ejemplos.

De modo análogo, un símbolo funcional binario, como es el símbolo $\underline{h}$ del lenguaje $\mathscr{K}$, deberá ir seguido de otros dos términos, para formar un nuevo término. Y así ocurre en efecto con los términos $\underline{hv_1v_1}$, $\underline{hv_1v_7}$ y $\underline{hdd}$, que aparecen en la 3ª línea de nuestra lista de ejemplos.

Y finalmente, hay que tener en cuenta también que, una vez formado un término, como pueda ser por ejemplo $\underline{gv_7}$, éste puede pasar a formar parte de otro término. Es por eso que $\underline{hv_1gv_7}$ (en la 4ª línea de nuestra lista), constituye todo él un término, porque empieza por un símbolo funcional binario, que es $\underline{h}$, seguido por dos términos, que son $\underline{v_1}$ y el propio $\underline{gv_7}$.

Queda claro que así sucesivamente se pueden ir formando términos cada vez más complicados, como por ejemplo $\underline{hv_7gv_7}$ y $\underline{hv_7hdd}$ (citados en último lugar en nuestra lista), y otros muchísimo más largos y complejos que estos.

§ 2.9. Definición (Subtérmino). Un *subtérmino* es una subcadena de un término que a su vez constituye ella misma un término.

Por ejemplo, siguiendo con el lenguaje precedente, son *subtérminos* del término $\underline{hv_7hdd}$ los términos $\underline{v_7}$, $\underline{d}$, $\underline{hdd}$, y por supuesto el propio $\underline{hv_7hdd}$. La cadena $\underline{v_7hd}$, en cambio, también es una subcadena de $\underline{hv_7hdd}$, pero no es un subtérmino porque ella misma no constituye un término.

§ 2.10. Definición (Grado de un término). Por su parte, llamamos *"grado"* de un término al número total de ocurrencias de constantes y símbolos funcionales en ese término. Los términos de grado 0, por consiguiente, serán las variables.

Por ejemplo, siguiendo con el lenguaje $\mathscr{K}$ de § 2.8 (p. 46), $\textit{grado}\,(\underline{v_7}) = 0$, $\textit{grado}\,(\underline{d}) = 1$, $\textit{grado}\,(\underline{gd}) = 2$, y $\textit{grado}\,(\underline{hv_7hdd}) = 4$.

§ 2.11. Definición (Peso). Dependiendo del tipo de símbolo, tenemos:

- El *peso* de una variable o de una constante es -1.
- El *peso* de un símbolo funcional o relacional n-ario es $n-1$.
- El *peso* de $\underline{\neg}$ es 0.
- El *peso* de $\underline{\rightarrow}$ y de $\underline{\forall}$ es 1.

Y por su parte, el *peso* de una cadena no es más que la suma de los pesos correspondientes a cada *ocurrencia* de un símbolo en esa cadena. Adviértase que esto se aplica a *cualquier cadena* de símbolos de un lenguaje de primer orden, no sólo a los términos.

Así, siguiendo con nuestro lenguaje de marras, tenemos: $peso\,(\underline{v_7}) = -1$, $peso\,(\underline{d}) = -1$, $peso\,(\underline{g}) = 1-1=0$, y $peso\,(\underline{h}) = 2-1=1$. Con lo que por ejemplo,

$$peso\,(\underline{hv_7gv_7}) = 1-1+0-1 = -1$$

§ 2.12. Práctica (Términos). Siguiendo con el lenguaje $\mathscr{K}$ de § 2.8, indicar 3 nuevos términos de este lenguaje, tan variados como sea posible. A continuación, calcular el grado y peso de cada uno, y delimitar exhaustivamente todos sus subtérminos, calculando asimismo el peso de cada uno de ellos.

§ 2.13. Lema (Peso de los términos). *Sea $\underline{t}$ un término cualquiera de un lenguaje de primer orden. Entonces:*

(a) $peso\,(\underline{t}) = -1$.

(b) Si $\underline{r}$ es cualquier subcadena inicial propia de $\underline{t}$, entonces $peso\,(\underline{r}) \geqslant 0$.

Prueba. Probamos los dos resultados simultáneamente, por inducción fuerte sobre el grado de $\underline{t}$. Para ello suponemos que cualquier término de su mismo lenguaje y que tenga grado menor que $\underline{t}$, cumple ya con el lema. Además, distinguimos y tratamos separadamente cada una de las tres formas distintas que puede tomar $\underline{t}$ de acuerdo con la definición de *término*.

En primer lugar, supongamos que $\underline{t}$ sea una variable $\underline{x}$. Entonces, por definición, $peso\,(\underline{x}) = -1$. Y al constituir $\underline{x}$ un único símbolo, no puede tener subcadenas propias. Así que las dos partes del lema se cumplen.

En segundo lugar, supongamos que $\underline{t}$ sea una constante $\underline{c}$. Entonces, al igual que antes, $peso\,(\underline{c}) = -1$, y $\underline{c}$ tampoco puede tener subcadenas propias. Así que las dos partes del lema se cumplen también.

Y en tercer lugar, supongamos que $\underline{t}$ tenga la forma

$$\underline{ft_1t_2\ldots t_n}$$

donde f es un símbolo funcional n-ario de $\mathscr{L}$ y $\underline{t_1}$, $\underline{t_2}$, ..., $\underline{t_n}$ son términos de $\mathscr{L}$.

Naturalmente, $grado\,(\underline{t_1}) < grado\,(\underline{t})$, ya que $\underline{t}$ contiene todas las ocurrencias de los símbolos funcionales que puedan aparecer en $\underline{t_1}$, y contiene además una ocurrencia de $\underline{f}$, la inicial, que no pertenece a $\underline{t_1}$. Por lo tanto podemos aplicar la hipótesis de inducción al término $\underline{t_1}$, y tenemos *peso* $(\underline{t_1}) = -1$.

Lo mismo ocurre con $\underline{t_2}$, $\ldots$, $\underline{t_n}$: el peso de todos estos términos es -1. Y como $\underline{f}$ es un símbolo funcional n-ario, tenemos *peso* $(\underline{f}) = n - 1$, con lo cual

$$\mathit{peso}\,(\underline{ft_1t_2\ldots t_n}) = (n-1)\overbrace{-1-1\ldots-1}^{n\text{ veces}} = (n-1) - n = -1$$

Y con ello hemos probado que $\underline{t}$ satisface la primera parte del lema.

Ahora sea $\underline{r}$ cualquier subcadena inicial propia de $\underline{t}$. Distinguimos aquí a su vez tres posibilidades. La primera es que $\underline{r}$ consista exclusivamente en el símbolo funcional $\underline{f}$. En este caso, como *peso* $(\underline{f}) = n - 1$, y por definición de *símbolo funcional* tenemos $n \geqslant 1$, podemos concluir que $n - 1 \geqslant 0$ y por tanto *peso* $(\underline{f}) \geqslant 0$.

La segunda posibilidad es que $\underline{r}$ comprenda el símbolo funcional $\underline{f}$ seguido de *algunos* de los términos $\underline{t_1}$, $\underline{t_2}$, $\ldots$, $\underline{t_n}$, pero no de todos: no pueden ser todos, porque entonces $\underline{r}$ no sería una subcadena *propia* de $\underline{t}$. En definitiva, en este caso $\underline{r}$ tendrá la forma

$$\underline{ft_1t_2\ldots t_m}$$

para algún m tal que $1 \leqslant m < n$.

Entonces, aplicando otra vez la hipótesis de inducción a cada uno de los términos $\underline{t_1}$, $\underline{t_2}$, $\ldots$, $\underline{t_m}$, tenemos

$$\mathit{peso}\,(\underline{ft_1t_2\ldots t_m}) = (n-1)\overbrace{-1-1\ldots-1}^{m\text{ veces}} = (n-1) - m$$

Ahora bien:

$$n - 1 - m \;=\; n - m - 1$$

Como n es estrictamente mayor que m, tenemos que $n-m \geqslant 1$, y por tanto $n-m-1 \geqslant 0$.

Y finalmente, la tercera posibilidad es que $\underline{r}$ consista en la cadena $\underline{ft_1t_2\ldots t_k}$ *más* una subcadena inicial propia del término $\underline{t_{k+1}}$. Es decir, que $\underline{r}$ tenga la forma

$$\underline{ft_1t_2\ldots t_kr'}$$

donde $\underline{r'}$ es una subcadena inicial propia de $\underline{t_{k+1}}$.

En este caso estamos considerando algún k tal que $0 \leqslant k < n$. De forma que, en particular, si $k = 0$ entonces el fragmento $\underline{r'}$ sería una subcadena inicial propia de $\underline{t_{0+1}}$, esto es, el fragmento $\underline{r'}$ sería una subcadena inicial propia de $\underline{t_1}$.

En dichas condiciones, aplicando la hipótesis de inducción a cada uno de los términos $\underline{t_1}$, $\underline{t_2}$, $\ldots$, $\underline{t_k}$, tenemos que

$$\mathit{peso}\,(\underline{ft_1t_2\ldots t_kr'}) \;=\; (n-1)\;\overbrace{-1-1\ldots-1}^{k\text{ veces}} + \mathit{peso}\,(r') \;=\; (n-1) - k + \mathit{peso}\,(r')$$

Como n es estrictamente mayor que k, tenemos otra vez que $(n-1)-k \geqslant 0$. Y por último, aplicando la hipótesis de inducción al término $\underline{t_{k+1}}$, al ser $\underline{r'}$ una subcadena propia de este término, podemos concluir que *peso* $(r') \geqslant 0$, con lo cual al sumarlo al resultado anterior el total seguirá siendo $\geqslant 0$.

§ 2.14. Observación (Escrutinio de los términos). Supongamos que $\underline{t}$ sea un término de un lenguaje de primer orden, y que en el interior de $\underline{t}$ aparezca un símbolo funcional n-ario $\underline{f}$. Entonces, para determinar cuál es el primer argumento de $\underline{f}$, basta con ir sumando los pesos de los símbolos que haya a la derecha de $\underline{f}$, sin contar a la propia $\underline{f}$, hasta que el resultado sea -1. Justo en ese punto, de acuerdo con el lema precedente, tendremos un término, y por la definición de *término* ése tendrá que ser a la fuerza el primer argumento de $\underline{f}$.

Para determinar el segundo argumento de $\underline{f}$ sumamos los pesos de los símbolos que encontremos a partir del primer argumento, y así sucesivamente.

Es claro que, en virtud del lema precedente, esta operación sólo puede dar un resultado, con lo que queda demostrado que la delimitación de los argumentos de un símbolo funcional, en el interior de un término, está unívocamente determinada.

Además, una vez completamente efectuado este análisis, resulta inmediato comprobar si los argumentos de los otros símbolos funcionales de $\underline{t}$ constituyen a su vez términos, y por consiguiente, si la cadena inicial era ella misma un término o no.

Fórmulas

§ 2.15. Definición (Fórmula). Decimos que una cadena $\underline{\alpha}$ de un lenguaje de primer orden $\mathscr{L}$ es una *"fórmula"* cuando cumple con alguna de estas cuatro condiciones:

1. $\underline{\alpha}$ es una cadena compuesta por un símbolo relacional n-ario $\underline{R}$ de $\mathscr{L}$ seguido de otras cadenas $\underline{t_1}, \underline{t_2}, \ldots, \underline{t_n}$, todas las cuales son términos de $\mathscr{L}$:
$$\underline{Rt_1t_2\ldots t_n}$$
2. $\underline{\alpha}$ es una cadena compuesta por una ocurrencia del símbolo de negación seguida de otra cadena $\underline{\beta}$, que es a su vez una *fórmula* de $\mathscr{L}$. Es decir, $\underline{\alpha}$ tiene la forma
$$\underline{\neg\beta}$$
(leído *"no $\underline{\beta}$"*).
3. $\underline{\alpha}$ es una cadena compuesta por una ocurrencia del símbolo condicional seguida de otras cadenas $\underline{\beta}$ y $\underline{\gamma}$, que a su vez son *fórmulas* de $\mathscr{L}$. Es decir, $\underline{\alpha}$ tiene la forma
$$\underline{\rightarrow\beta\gamma}$$
(leído *"si $\underline{\beta}$ entonces $\underline{\gamma}$"*).
4. $\underline{\alpha}$ es una cadena compuesta por una ocurrencia del símbolo de cuantificación universal seguida por una variable $\underline{x}$ de $\mathscr{L}$ y por otra cadena $\underline{\beta}$, que a su vez es una *fórmula* de $\mathscr{L}$. Es decir:
$$\underline{\forall x\beta}$$
(leído *"para todo $\underline{x}$, $\underline{\beta}$"*).

La definición de "fórmula" también procede por inducción fuerte sobre la longitud de la cadena, tomando una cadena arbitraria del lenguaje $\mathscr{L}$, la cadena $\underline{\alpha}$, y especificando las condiciones bajo las cuales $\underline{\alpha}$ es una fórmula, dada la suposición de que ya sabemos determinar si son fórmulas o no las cadenas de longitud menor que $\underline{\alpha}$. En particular, nótese que las cadenas $\underline{\beta}$ y $\underline{\gamma}$ de las cláusulas (2), (3) y (4), tienen todas ellas longitud forzosamente menor que $\underline{\alpha}$.

Nótese que, en particular, la cláusula (1) se aplica también al caso del símbolo de igualdad, $\underline{=}$, que como hemos dicho constituye un símbolo relacional binario. Por consiguiente, si $\underline{t_1}$ y $\underline{t_2}$ son términos de $\mathscr{L}$, la correspondiente cadena $\underline{= t_1 t_2}$ será una fórmula de $\mathscr{L}$, de acuerdo con la definición que se acaba de enunciar.

Y ahora, volviendo por ejemplo a nuestro lenguaje $\mathscr{K}$ de § 2.8, podemos comprobar que las siguientes cadenas son casos claros de fórmulas de dicho lenguaje:

$$
\begin{array}{rrr}
\underline{Pv_7} & \underline{Pd} & \underline{= v_7 d} \\
\underline{\neg Pv_7} & \underline{\neg = v_7 d} & \underline{\neg\neg = v_7 d} \\
\underline{Pgv_1} & \underline{= gv_1 v_7} & \underline{= gv_1 hv_7 hdd} \\
\underline{\rightarrow Pv_7 Pd} & \underline{\neg \rightarrow Pv_7 Pd} & \underline{\rightarrow Pv_7 \neg\neg = v_7 d} \\
\underline{\forall v_1 Pv_1} & \underline{\forall v_7 Pv_7} & \underline{\neg \forall v_7 \neg Pv_7} \\
\underline{\forall v_7 = v_1 v_1} & \underline{\forall v_7 \rightarrow Pv_7 Pv_7} & \underline{\forall v_1 \forall v_2 \forall v_3 \rightarrow Pv_7 \neg\neg = v_7 d}
\end{array}
$$

§ 2.16. Definición (Tipos de fórmulas). En el caso en que una fórmula tenga la forma $\underline{Rt_1 t_2 \ldots t_n}$ decimos que es una *"fórmula atómica"*. Si, en particular, tiene la forma $\underline{= t_1 t_2}$ (leído *"$\underline{t_1}$ igual a $\underline{t_2}$"*), entonces decimos también que es una *"fórmula de igualdad"*.

En el caso en que una fórmula tenga la forma $\underline{\neg\beta}$, decimos que la ocurrencia inicial de $\underline{\neg}$ es su *"operador principal"*, y que la *fórmula* $\underline{\beta}$ es el *"alcance"* de este operador. Decimos también que $\underline{\neg\beta}$ es una *"fórmula de negación"*, y que $\underline{\beta}$ es la *"fórmula negada"* en $\underline{\neg\beta}$.

En el caso en que una fórmula tenga la forma $\underline{\rightarrow \beta\gamma}$, decimos que la ocurrencia inicial de $\underline{\rightarrow}$ es su *"operador principal"*, y que el *"alcance"* de este operador *abarca* a las *fórmulas* $\underline{\beta}$ y $\underline{\gamma}$. En este caso decimos también que $\underline{\rightarrow \beta\gamma}$ es una *"fórmula condicional"*, y que $\underline{\beta}$ y $\underline{\gamma}$ son, respectivamente, la *"fórmula de antecedente"* y la *"fórmula de consecuente"* de dicha fórmula.

Y finalmente, en el caso en que una fórmula tenga la forma $\underline{\forall x\beta}$, decimos que la ocurrencia inicial de $\underline{\forall}$ es su *"operador principal"*, y que este operador tiene a $\underline{x}$ como "variable de cuantificación" y a la fórmula $\underline{\beta}$ como "alcance". Y decimos también que la fórmula $\underline{\forall x\beta}$ en su conjunto es una *"fórmula de cuantificación universal"*, y que $\underline{\beta}$ es la *"fórmula cuantificada universalmente"* en ella.

§ 2.17. Definición (Subfórmula). Una *subfórmula* es una subcadena de una fórmula que a su vez constituye ella misma una fórmula.

Por ejemplo, siguiendo con el lenguaje precedente, son *subfórmulas* de

$$\underline{\forall v_1 \forall v_2 \forall v_3 \rightarrow Pv_7 \neg\neg = v_7 d}$$

las fórmulas

$$
\begin{array}{lrr}
\underline{Pv_7} & \underline{= v_7 d} & \underline{\rightarrow Pv_7 \neg\neg = v_7 d} \\
 & \underline{\neg = v_7 d} & \underline{\forall v_3 \rightarrow Pv_7 \neg\neg = v_7 d} \\
 & \underline{\neg\neg = v_7 d} & \underline{\forall v_2 \forall v_3 \rightarrow Pv_7 \neg\neg = v_7 d}
\end{array}
$$

y por supuesto la propia $\underline{\forall v_1 \forall v_2 \forall v_3 \to P v_7 \neg\neg = v_7 d}$. La cadena $\underline{v_2 \forall v_3 \to P}$, en cambio, también es una subcadena suya, pero no es una subfórmula porque ella misma no constituye una fórmula.

§ 2.18. Definición (Grado de una fórmula). A su vez, el *grado* de una fórmula es el número total de ocurrencias de conectivas y del símbolo de cuantificación universal en esa fórmula. Las fórmulas de grado 0, por lo tanto, serán las fórmulas atómicas.

Por ejemplo, siguiendo con nuestro manido lenguaje $\mathscr{K}$ de § 2.8,

$$\begin{array}{rr} \mathit{grado}\,(\underline{= v_7 d}) = 0 & \mathit{grado}\,(\underline{\neg\neg = v_7 d}) = 2 \\ \mathit{grado}\,(\underline{\neg = v_7 d}) = 1 & \mathit{grado}\,(\underline{\to P v_7 \neg\neg = v_7 d}) = 3 \end{array}$$

§ 2.19. Práctica (Fórmulas). Siguiendo con el mismo lenguaje de primer orden, indicar 3 nuevas fórmulas de ese lenguaje, tan variadas como sea posible. Determinar sus operadores principales, los alcances de estos operadores, y, donde proceda, la división entre fórmula de antecedente y fórmula de consecuente, y la posición de la variable de cuantificación. Finalmente, calcular el grado de cada una.

§ 2.20. Teorema (Peso de las fórmulas). *Sea $\underline{\alpha}$ una fórmula cualquiera de un lenguaje de primer orden. Entonces:*

(a) peso $(\underline{\alpha}) = -1$.

(b) Si $\underline{r}$ es cualquier subcadena inicial propia de $\underline{\alpha}$, entonces peso $(\underline{r}) \geqslant 0$.

Prueba. Muy similar a la del lema sobre el peso de los términos, queda como ejercicio que a continuación se propone.

§ 2.21. Práctica* (Peso de las fórmulas). Probar el teorema sobre el peso de las fórmulas que se acaba de enunciar.

§ 2.22. Observación (Escrutinio de las fórmulas). Es claro que, de acuerdo con el teorema precedente, para determinar cuál es el alcance de una ocurrencia de $\underline{\neg}$ en una fórmula, basta con ir sumando los pesos de los símbolos que haya a la derecha de esa ocurrencia hasta que el resultado sea -1. En el caso de $\underline{\to}$, por este procedimiento se delimita la fórmula de antecedente, y repitiéndolo con los símbolos que hay a la derecha de la fórmula delimitada, se determina a su vez la fórmula de consecuente. Y en el caso de $\underline{\forall}$, debemos encontrar una variable a la derecha de $\underline{\forall}$, y empezar a contar a la derecha de esa variable.

Obviamente estas operaciones sólo pueden dar un resultado en cada caso, con lo que queda demostrado que la delimitación de alcances, y la separación entre las fórmulas de antecedente y de consecuente en el interior de las fórmulas condicionales, están unívocamente determinadas.

Además, una vez completamente efectuado este análisis resulta inmediato comprobar si las subcadenas correspondientes son a su vez fórmulas de acuerdo con la definición, y por consiguiente, si la cadena inicial era ella misma una fórmula o no.

§ 2.23. Convención (Algunas abreviaturas para fórmulas).

1. Siendo $\underline{r}$ y $\underline{t}$ términos de algún lenguaje de primer orden, entonces:
   - A una fórmula del tipo $\underline{= rt}$ la denotaremos también poniendo "$\underline{(r = t)}$".
   - A una fórmula del tipo $\underline{\neg\,(r = t)}$ la denotaremos también poniendo "$\underline{(r \neq t)}$".
2. Siendo $\underline{\beta}$ y $\underline{\gamma}$ fórmulas cualesquiera de algún lenguaje de primer orden, entonces:
   - A una fórmula del tipo $\underline{\to \beta\gamma}$ la denotaremos también poniendo "$\underline{(\beta \to \gamma)}$".
   - A una fórmula del tipo $\underline{(\neg\beta \to \gamma)}$ la denotaremos también poniendo "$\underline{(\beta \vee \gamma)}$".
   - A una fórmula del tipo $\underline{\neg(\beta \to \neg\gamma)}$ la denotaremos también poniendo "$\underline{(\beta \wedge \gamma)}$".
   - A una fórmula del tipo $\underline{((\beta \to \gamma) \wedge (\gamma \to \beta))}$ la denotaremos también poniendo "$\underline{(\beta \leftrightarrow \gamma)}$".
3. Finalmente, siendo $\underline{x}$ una variable cualquiera de un lenguaje de primer orden, y $\underline{\beta}$ una fórmula de ese mismo lenguaje, entonces
   - A una fórmula del tipo $\underline{\neg\forall x \neg\beta}$ la denotaremos también poniendo "$\underline{\exists x \beta}$".

A una fórmula como $\underline{(r \neq t)}$ (leído *"r distinto de t"*) la llamaremos también *"fórmula de desigualdad"*.

A una fórmula como $\underline{(\beta \vee \gamma)}$ (leído *"$\underline{\beta}$ o $\underline{\gamma}$"*) la llamamos también *"fórmula de disyunción"*, y diremos que las fórmulas $\underline{\beta}$ y $\underline{\gamma}$ son, respectivamente, su *"fórmula de primer disyunto"* y su *"fórmula de segundo disyunto"*.

A una fórmula como $\underline{\beta \wedge \gamma}$ (leído *"$\underline{\beta}$ y $\underline{\gamma}$"*) la llamamos también *"fórmula de conjunción"*, y diremos que las fórmulas $\underline{\beta}$ y $\underline{\gamma}$ son, respectivamente, su *"fórmula de primer conyunto"* y su *"fórmula de segundo conyunto"*.

A una fórmula como $\underline{(\beta \leftrightarrow \gamma)}$ (leído *"$\underline{\beta}$ si y sólo si $\underline{\gamma}$"*) la llamamos también *"fórmula bicondicional"*.

Y a una fórmula como $\underline{\exists x \beta}$ (leído *"existe un $\underline{x}$ tal que $\underline{\beta}$"*) la llamaremos también *"fórmula de cuantificación existencial"*, y diremos que $\underline{\beta}$ es la *"fórmula cuantificada existencialmente"* en ella.

Por ejemplo, una vez más con nuestro lenguaje $\mathscr{K}$ de § 2.8,

$$\begin{array}{rcl}
\underline{(Pv_7 \to Pd)} & \text{es la misma fórmula que} & \underline{\to Pv_7 Pd} \\
\underline{(Pd \wedge Pv_7)} & " & \underline{\neg \to Pd \neg Pv_7} \\
\underline{(\exists v_7 Pv_7)} & " & \underline{\neg\forall v_7 \neg Pv_7} \\
\underline{\forall v_1 \forall v_2 \forall v_3 (Pv_7 \to \neg(v_7 \neq d))} & " & \underline{\forall v_1 \forall v_2 \forall v_3 \to Pv_7 \neg\neg = v_7 d}
\end{array}$$

§ 2.24. Convención (Más abreviaturas para fórmulas). Si $\underline{\beta_1}$, $\underline{\beta_2}$, ..., $\underline{\beta_n}$, para algún número natural $n \geqslant 3$, son fórmulas de un lenguaje de primer orden, entonces a una fórmula del tipo

$$\underline{(\beta_1 \vee (\beta_2 \vee (\ldots \vee \beta_n)\ldots))}$$

la denotaremos también poniendo

$$\text{“}\underline{(\beta_1 \vee \beta_2 \vee \ldots \vee \beta_n)}\text{”}$$

Y a una fórmula del tipo

$$\underline{(\beta_1 \wedge (\beta_2 \wedge (\ldots \wedge \beta_n) \ldots))}$$

la denotaremos también poniendo

$$\text{“}\underline{(\beta_1 \wedge \beta_2 \wedge \ldots \wedge \beta_n)}\text{”}$$

Además, si $\underline{x_1}$, $\underline{x_2}$, ..., $\underline{x_n}$, para algún número natural $n \geqslant 2$, son variables cualesquiera de un lenguaje de primer orden $\mathscr{L}$, y $\underline{\beta}$ es una fórmula de ese lenguaje, entonces a una fórmula del tipo

$$\underline{\forall x_1 \forall x_2 \ldots \forall x_n \beta}$$

la denominaremos también poniendo

$$\text{“}\underline{\forall x_1 x_2 \ldots x_n \beta}\text{”}$$

Y a una fórmula del tipo

$$\underline{\exists x_1 \exists x_2 \ldots \exists x_n \beta}$$

la denominaremos también poniendo

$$\text{“}\underline{\exists x_1 x_2 \ldots x_n \beta}\text{”}$$

Finalmente, donde quiera que una fórmula $\underline{(\alpha)}$ de un lenguaje de primer orden se nombre sola, es decir, no como subcadena de otra fórmula, convenimos en que podemos denotarla como “$\underline{\alpha}$” sin más. Es decir: que nos permitimos suprimir los paréntesis más exteriores cada vez que mencionamos una determinada fórmula.

Por ejemplo, volviendo a nuestro lenguaje $\mathscr{K}$ de siempre,

$$\begin{array}{rcl} \underline{Pv_7 \to Pd} & \text{es la misma fórmula que} & \underline{(Pv_7 \to Pd)} \\ \underline{\forall v_1 v_2 v_3 (Pv_7 \to \neg(v_7 \neq d))} & \text{”} & \underline{\forall v_1 \forall v_2 \forall v_3 (Pv_7 \to \neg(v_7 \neq d))} \end{array}$$

§ 2.25. Observación (Signos introducidos en las abreviaturas). Algunas de las abreviaturas anteriores incluyen el uso de paréntesis, de los que hasta ahora no habíamos tenido necesidad, así como de otros signos especiales, como los signos para la conjunción, disyunción, etc. Sin embargo, se ha de tener muy claro que los lenguajes de primer orden tal y como los hemos definido nosotros, no contienen ninguno de esos signos adicionales. Dichos signos son meras herramientas que utilizamos nosotros para referirnos a los términos y fórmulas de los lenguajes de primer orden. Y pertenecen, por lo tanto, a nuestro metalenguaje, pero no a nuestro lenguaje objeto.

Esta forma de notación, que hace innecesario el uso de paréntesis en el lenguaje objeto, se basa esencialmente en el hecho de que el símbolo condicional aparezca precediendo a antecedente y consecuente, en lugar de ir en medio. Por eso se le llama *“notación prefija”*, así como también *“notación polaca”*, en honor al lógico polaco Jan Lukasiewicz, que la inventó.

En otros tratamientos de la lógica de primer orden distintos al del presente Curso, es posible que la definición de *lenguaje de primer orden* escogida sí incluya los paréntesis, u otros signos especiales, entre los símbolos originales del alfabeto, o bien que deje fuera a alguno de los nuestros. No importa: el resultado final en cuanto a capacidad expresiva de estos lenguajes debe ser esencialmente equivalente.

§ 2.26. Práctica (Abreviaturas para fórmulas). Volver a mencionar las 3 fórmulas propuestas en § 2.19 (p. 52), pero utilizando las abreviaturas introducidas donde sea posible.

Modelos

§ 2.27. Definición (Modelo). Una *interpretación* o *modelo* $\mathscr{M}$ para un lenguaje de primer orden $\mathscr{L}$ es una entidad compuesta por cuatro ingredientes distintos, que son:

1. Un conjunto no vacío cualquiera U, llamado *"universo"* de $\mathscr{M}$.
2. Una función que asigna, a cada constante $\underline{c}$ de $\mathscr{L}$, un elemento cualquiera de U, que representamos como "$\underline{c}^{\mathscr{M}}$" (leído *"$\underline{c}$ bajo $\mathscr{M}$"*), y que constituirá el *valor* de $\underline{c}$ en $\mathscr{M}$.
3. Una función que asigna, a cada símbolo funcional n-ario $\underline{f}$ de $\mathscr{L}$, una función n-aria que representamos como "$\underline{f}^{\mathscr{M}}$" (leído *"$\underline{f}$ bajo $\mathscr{M}$"*), y que constituirá el *valor* de $\underline{f}$ en $\mathscr{M}$. Esta función $\underline{f}^{\mathscr{M}}$ ha de tener como dominio a U, y ha de dar como valores siempre elementos de U.
4. Una función que asigna, a cada símbolo relacional n-ario $\underline{R}$ de $\mathscr{L}$ distinto del símbolo de igualdad, una relación n-aria entre elementos de U, que denotamos como "$\underline{R}^{\mathscr{M}}$" (leído *"$\underline{R}$ bajo $\mathscr{M}$"*), y que constituirá el *valor* de $\underline{R}$ en $\mathscr{M}$.

Por ejemplo, el modelo $\mathscr{S}$ que a continuación se describe es un caso de modelo para nuestro heroico lenguaje $\mathscr{K}$ de § 2.8 (p. 46):

- Como universo, $\mathscr{S}$ posee un conjunto W compuesto por tres individuos distintos, que vamos a poner que sean Don Quijote, Sancho Panza y Dulcinea del Toboso, representados por sus iniciales, las letras *"DQ"*, *"SP"* y *"DT"* respectivamente:
$$W = \{ DQ, SP, DT \}$$
- El valor de la constante $\underline{d}$ en $\mathscr{S}$ será el individuo DT, esto es, Dulcinea:
$$\underline{d}^{\mathscr{S}} = DT$$
- El valor del símbolo funcional $\underline{g}$ en $\mathscr{S}$ es aquella función monaria i tal que:
$$i(DQ) = SP \qquad i(SP) = DT \qquad i(DT) = DQ$$

- El valor del símbolo funcional $\underline{h}$ en $\mathscr{S}$ es aquella función binaria j que siempre da DQ como resultado. Es decir, tal que:

$$j(DQ, DQ) = DQ \qquad j(DQ, SP) = DQ \qquad j(SP, DQ) = DQ \qquad \text{etc.}$$

- El valor del símbolo predicativo $\underline{P}$ es aquella propiedad Q que cumplen únicamente los objetos DQ y SP, esto es, Don Quijote y Sancho (podría corresponder, por ejemplo, a la propiedad de *ser varón*):

$$Q(DQ) \qquad Q(SP) \qquad \backsim Q(DT)$$

§ 2.28. Práctica (Modelos). Describir un nuevo ejemplo de modelo para el mismo lenguaje, cuyo universo tenga exactamente cuatro elementos distintos.

§ 2.29. Definición (Evaluación). Una *evaluación* ε para un lenguaje de primer orden $\mathscr{L}$ es el resultado de añadir a un modelo $\mathscr{M}$ para $\mathscr{L}$ una función que asigne, a cada variable $\underline{x}$ de $\mathscr{L}$, un elemento cualquiera del universo, que denotamos como "$\underline{x}^{\,\varepsilon}$" (leído *"$\underline{x}$ bajo ε"*), y que constituirá el *valor* de $\underline{x}$ en ε.

En definitiva, ε es una entidad compuesta por *cinco* ingredientes distintos, cuyos cuatro primeros coinciden con los de $\mathscr{M}$, y el quinto es la función que asigna valores a las variables y que acabamos de describir. Por ello decimos en estos casos que la evaluación ε está *"basada"* en el modelo $\mathscr{M}$.

Así pues, ε tendrá un universo, que será el mismo universo del modelo $\mathscr{M}$. Además, si $\underline{c}$ es una constante de $\mathscr{L}$, entonces $\underline{c}$ tendrá en ε un *valor*, $\underline{c}^{\,\varepsilon}$, que será exactamente el mismo que tiene en $\mathscr{M}$, es decir:

$$\underline{c}^{\,\varepsilon} = \underline{c}^{\,\mathscr{M}}$$

Y otro tanto sucede con cualquier otra constante de $\mathscr{L}$, así como con todos sus símbolos funcionales, y con todos sus símbolos relacionales distintos al de igualdad.

En resumidas cuentas: si un modelo para $\mathscr{L}$ incorpora una especie de "diccionario", valga la comparación, mediante el cual se atribuye significado a cada uno de sus símbolos peculiares, la evaluación viene a ser algo así como un posible "suplemento" asociado a dicho diccionario, en el cual se fijan, además, los significados de las variables, de una determinada manera.

Por ejemplo, siguiendo con el modelo $\mathscr{S}$ definido en § 2.27, una evaluación basada en $\mathscr{S}$ sería, pongamos por caso, aquella evaluación δ que asignase a todas las variables de $\mathscr{K}$ el individuo SP, esto es, Sancho Panza. Es decir:

$$\underline{v_1}^{\,\delta} = SP \qquad \underline{v_2}^{\,\delta} = SP \qquad \underline{v_3}^{\,\delta} = SP \qquad \underline{v_4}^{\,\delta} = SP \qquad \ldots$$

O lo que es lo mismo:

$$\underline{x}^{\,\delta} = SP \qquad \text{para cualquier variable } \underline{x} \text{ de } \mathscr{K}$$

Otro ejemplo de evaluación basada en $\mathscr{S}$ lo constituiría aquella evaluación δ' que asignase a todas las variables de $\mathscr{K}$ el individuo SP, excepto a la variable $\underline{v_1}$, a la cual asignase el individuo DT, o sea, Dulcinea. Es decir:

$$\underline{v_1}^{\,\delta'} = DT \qquad \underline{v_2}^{\,\delta'} = SP \qquad \underline{v_3}^{\,\delta'} = SP \qquad \underline{v_4}^{\,\delta'} = SP \qquad \ldots$$

§ 2.30. Definición (Cardinalidad de modelos y evaluaciones). La *cardinalidad de un modelo* para un lenguaje de primer orden será simplemente la cardinalidad del conjunto que tenga como universo. Y la *cardinalidad de una evaluación* será a su vez la cardinalidad del modelo en que esté basada.

§ 2.31. Definición (Reasignación). Ahora sean:

- ε una evaluación cualquiera para un lenguaje de primer orden $\mathscr{L}$
- $\underline{x}$ una variable de $\mathscr{L}$
- a un elemento del universo de ε

Entonces, al *reasignar* en ε la variable $\underline{x}$ con el objeto a, se obtiene una evaluación, $\varepsilon\,[\underline{x}/a]$, que asigna a $\underline{x}$ el objeto a y que coincide en todo lo demás con la evaluación ε.

Por tanto, el valor de $\underline{x}$ en $\varepsilon\,[\underline{x}/a]$ es a:

$$\underline{x}^{\,\varepsilon\,[\underline{x}/a]} = a$$

Y en todo lo demás, $\varepsilon\,[\underline{x}/a]$ se comporta exactamente igual que ε, es decir: $\varepsilon\,[\underline{x}/a]$ tiene el mismo universo que ε, y el valor en $\varepsilon\,[\underline{x}/a]$, de cualquier variable distinta de $\underline{x}$, es el mismo que tenía en ε.

Por ejemplo, siguiendo con nuestras evaluaciones anteriores, es obvio que,

$$\delta\,[\underline{v_1}/DT] = \delta'$$

mientras que, por su parte, si efectuamos la reasignación $\delta\,[\underline{v_1}/SP]$, ésta quedará sin efecto, ya que el valor de $\underline{v_1}$ en δ era ya de hecho SP. Por consiguiente:

$$\delta\,[\underline{v_1}/SP] = \delta$$

§ 2.32. Práctica (Evaluaciones). Proponer dos ejemplos distintos de evaluaciones, basadas en el modelo $\mathscr{S}$ de § 2.27, y efectuar una reasignación en cada una, describiendo cuál es el resultado.

§ 2.33. Definición (Valores de los términos). Sea $\mathscr{L}$ un lenguaje de primer orden y ε una evaluación para $\mathscr{L}$. De acuerdo con las definiciones de *modelo* y de *evaluación*, todas las variables y constantes de $\mathscr{L}$ reciben un valor en ε. Pues bien, la presente definición procede por inducción fuerte sobre el grado de un término cualquiera $\underline{t}$ de $\mathscr{L}$, adjudicándole un *valor*, $\underline{t}^{\,\varepsilon}$, en la evaluación ε.

Para ello suponemos que si un término de $\mathscr{L}$ tiene grado menor que $\underline{t}$, entonces ya sabemos el *valor* que le corresponde en dicha evaluación. Además, distinguimos y tratamos separadamente cada una de las tres formas que puede tomar $\underline{t}$ de acuerdo con la definición de *término*.

1. Supongamos en primer lugar que $\underline{t}$ sea una variable $\underline{x}$. Entonces, como se acaba de señalar, $\underline{x}$ tiene ya un valor en la evaluación ε.

2. A continuación, supongamos que $\underline{t}$ sea una constante. Pero también entonces tiene ya un valor en la evaluación ε, y no hay nada que definir.

3. Por último, supongamos que $\underline{t}$ tenga la forma

$$\underline{ft_1t_2\ldots t_n}$$

donde $\underline{f}$ es un símbolo funcional n-ario de $\mathscr{L}$ y $\underline{t_1}$, $\underline{t_2}$, …, $\underline{t_n}$ son términos de $\mathscr{L}$. Entonces el *valor*

$$\underline{ft_1t_2\ldots t_n}^{\,\varepsilon}$$

del término $\underline{ft_1t_2\ldots t_n}$ en la evaluación ε, es aquel elemento del universo de ε que resulta de aplicar la función $\underline{f}^{\,\varepsilon}$ a aquellos elementos del universo que constituyen, respectivamente, los valores de $\underline{t_1}$, $\underline{t_2}$, …, $\underline{t_n}$ en ε. Es decir:

$$\underline{ft_1t_2\ldots t_n}^{\,\varepsilon} = \underline{f}^{\,\varepsilon}\,(\,\underline{t_1}^{\,\varepsilon},\underline{t_2}^{\,\varepsilon},\ldots,\underline{t_n}^{\,\varepsilon}\,)$$

Por ejemplo, siguiendo con nuestra evaluación δ de § 2.29, tenemos, entre otros:

$$\begin{array}{rr} \underline{v_1}^{\,\delta} = SP & \underline{v_7}^{\,\delta} = SP \\ \underline{d}^{\,\delta} = DT & \underline{gv_7}^{\,\delta} = i(SP) = DT \\ \underline{gd}^{\,\delta} = i(DT) = DQ & \underline{hv_1gv_7}^{\,\delta} = j(SP,DT) = DQ \end{array}$$

§ 2.34. Práctica (Valores de los términos). Calcular los valores de estos mismos términos, bajo dos de las evaluaciones propuestas en la Práctica anterior.

§ 2.35. Definición (Valores de las fórmulas). Sea $\mathscr{L}$ un lenguaje de primer orden, y ε una evaluación para $\mathscr{L}$. La presente definición procede por inducción fuerte sobre el grado de una fórmula cualquiera $\underline{\alpha}$, adjudicándole un *valor*, $\underline{\alpha}^{\,\varepsilon}$, en la evaluación ε. Para ello suponemos que si una fórmula cualquiera de su mismo lenguaje tiene grado menor que $\underline{\alpha}$, entonces ya sabemos el *valor* que le corresponde en dicha evaluación. Además, distinguimos y tratamos separadamente cada una de las cuatro formas que puede tomar $\underline{\alpha}$ de acuerdo con la definición de *fórmula.*

En este caso sólo habrá dos valores posibles que las fórmulas puedan recibir: los dos valores de verdad, *verdadero* ($\mathbb{V}$) y *falso* ($\mathbb{F}$).

1. Supongamos en primer lugar que $\underline{\alpha}$ sea una fórmula de igualdad

$$\underline{r = t}$$

donde $\underline{r}$ y $\underline{t}$ son términos de $\mathscr{L}$. Entonces el *valor* de $\underline{\alpha}$ en ε será $\mathbb{V}$ cuando ε asigne a los términos $\underline{r}$ y $\underline{t}$ el mismo elemento del universo. Y $\mathbb{F}$ cuando les asigne elementos distintos. Es decir:

$$\underline{r = t}^{\,\varepsilon} = \begin{cases} \mathbb{V} & \text{si } \underline{r}^{\,\varepsilon} = \underline{t}^{\,\varepsilon} \\ \\ \mathbb{F} & \text{si } \underline{r}^{\,\varepsilon} \neq \underline{t}^{\,\varepsilon} \end{cases}$$

2. A continuación, supongamos que $\underline{\alpha}$ es una fórmula atómica

$$\underline{Rt_1t_2\dots t_n}$$

donde $\underline{R}$ es un símbolo relacional n-ario de $\mathscr{L}$, distinto del de igualdad, y $\underline{t_1}$, $\underline{t_2}$, ..., $\underline{t_n}$ son términos de $\mathscr{L}$. Entonces el *valor* de $\underline{\alpha}$ en ε será $\mathbb{V}$ cuando los elementos del universo que constituyan los valores de $\underline{t_1}$, $\underline{t_2}$, ..., $\underline{t_n}$ en ε cumplan, por ese orden, con la relación $\underline{R}^{\varepsilon}$. Y $\mathbb{F}$ en caso contrario. Es decir:

$$\underline{Rt_1t_2\dots t_n}^{\varepsilon} = \begin{cases} \mathbb{V} & \text{si } \underline{R}^{\varepsilon}(\underline{t_1}^{\varepsilon}, \underline{t_2}^{\varepsilon}, \dots, \underline{t_n}^{\varepsilon}) \\ \\ \mathbb{F} & \text{si} \sim \underline{R}^{\varepsilon}(\underline{t_1}^{\varepsilon}, \underline{t_2}^{\varepsilon}, \dots, \underline{t_n}^{\varepsilon}) \end{cases}$$

3. A continuación, supongamos que $\underline{\alpha}$ sea una negación

$$\underline{\neg\beta}$$

donde $\underline{\beta}$ es alguna otra fórmula de $\mathscr{L}$. Entonces el *valor* de $\underline{\alpha}$ en ε será $\mathbb{V}$ cuando el *valor* de $\underline{\beta}$ sea $\mathbb{F}$, y viceversa. Es decir:

$$\underline{\neg\beta}^{\varepsilon} = \begin{cases} \mathbb{V} & \text{si } \underline{\beta}^{\varepsilon} = \mathbb{F} \\ \\ \mathbb{F} & \text{si } \underline{\beta}^{\varepsilon} = \mathbb{V} \end{cases}$$

4. A continuación, supongamos que $\underline{\alpha}$ sea una fórmula condicional

$$\underline{\beta \to \gamma}$$

donde $\underline{\beta}$ y $\underline{\gamma}$ son fórmulas de $\mathscr{L}$. Entonces el *valor* de $\underline{\alpha}$ en ε será $\mathbb{F}$ cuando el *valor* de β sea $\mathbb{V}$ y el de γ sea $\mathbb{F}$. Y su valor será $\mathbb{V}$ en cualquier otro caso. O sea:

$$\underline{\beta \to \gamma}^{\varepsilon} = \begin{cases} \mathbb{F} & \text{si } \underline{\beta}^{\varepsilon} = \mathbb{V} \text{ y } \underline{\gamma}^{\varepsilon} = \mathbb{F} \\ \\ \mathbb{V} & \text{en cualquier otro caso} \end{cases}$$

En otras palabras: para que $\underline{\beta \to \gamma}^{\varepsilon}$ sea $\mathbb{V}$, tiene que suceder que *si* $\underline{\beta}^{\varepsilon}$ es $\mathbb{V}$, *entonces* $\underline{\gamma}^{\varepsilon}$ ha de ser también $\mathbb{V}$.

5. Y por último, supongamos que $\underline{\alpha}$ sea una cuantificación universal

$$\underline{\forall x\beta}$$

donde $\underline{x}$ es una variable, y $\underline{\beta}$ alguna otra fórmula, ambas de $\mathscr{L}$. Entonces el valor de $\underline{\alpha}$ en ε será *verdadero* cuando para cualquier elemento a del universo U de la evaluación, la evaluación reasignada $\varepsilon[\underline{x}/a]$ haga $\mathbb{V}$ a la fórmula $\underline{\beta}$. Y el valor de $\underline{\alpha}$ en ε será *falso* en caso contrario. Es decir:

$$\underline{\forall x\beta}^{\varepsilon} = \begin{cases} \mathbb{V} & \text{si para todo } a \in U \text{ tenemos } \underline{\beta}^{\varepsilon[\underline{x}/a]} = \mathbb{V} \\ \\ \mathbb{F} & \text{si existe algún } a \in U \text{ para el cual } \underline{\beta}^{\varepsilon[\underline{x}/a]} = \mathbb{F} \end{cases}$$

§ 2.36. Observación (Valores de las fórmulas). La definición precedente constituye la *"definición de verdad"* para nuestros lenguajes formales, y es de capital importancia.

Según la cláusula de la cuantificación, y dicho de una manera informal, la fórmula $\underline{\forall x \beta}$ será verdadera bajo ε cuando cualquier elemento del universo, de ser asignado a la variable $\underline{x}$, haría verdadera a la fórmula $\underline{\beta}$ bajo ε. El valor que de hecho ε asigna a la variable $\underline{x}$ no importa, como tampoco importa el valor que la fórmula $\underline{\beta}$ tenga bajo la evaluación ε. Lo que importa es que para cualquier elemento a del universo, al llevar a cabo la reasignación de ε aplicando a $\underline{x}$ ese elemento a, obtengamos el efecto de que la fórmula $\underline{\beta}$ sea verdadera bajo la evaluación reasignada resultante, $\varepsilon\,[\underline{x}/a]$.

En resumen: que $\underline{\forall x \beta}$ no es interpretada como afirmado que $\underline{\beta}$ es de hecho verdadera bajo la evaluación en cuestión, sino como afirmando que *cualquier elemento* del universo de esa evaluación la haría verdadera, si se le colocara como valor de la variable de cuantificación $\underline{x}$.

Y para el caso en que $\underline{\forall x \beta}$ resulte *falsa*, pues ocurrirá exactamente lo contrario: es decir, que tendrá que haber *al menos un* elemento a del universo que tenga el efecto de hacer *falsa* a $\underline{\beta}$ bajo $\varepsilon\,[\underline{x}/a]$.

Para poner algunos ejemplos, volveremos a nuestra conocida evaluación δ de § 2.29 (p. 56). Recordemos primero que

$$\underline{v_1}^{\delta} = SP \qquad \underline{v_2}^{\delta} = SP \qquad \underline{v_3}^{\delta} = SP \qquad \underline{v_4}^{\delta} = SP \qquad \ldots \qquad \underline{v_7}^{\delta} = SP \qquad \ldots$$

Y que, al estar δ basada en el modelo $\mathscr{S}$, tenemos:

$$\underline{d}^{\delta} = DT$$

y $\underline{P}^{\delta}$ es aquella propiedad Q tal que:

$$Q(DQ) \qquad\qquad Q(SP) \qquad\qquad \backsim Q(DT)$$

Por consiguiente, evidentemente,

$$\begin{array}{rr} \underline{Pv_7}^{\delta} = \mathbb{V} & \underline{Pd}^{\delta} = \mathbb{F} \\ \underline{\neg Pv_7}^{\delta} = \mathbb{F} & \underline{v_7 = d}^{\delta} = \mathbb{F} \\ \underline{Pd \to Pv_7}^{\delta} = \mathbb{V} & \underline{\neg(Pd \to Pv_7)}^{\delta} = \mathbb{F} \end{array}$$

Así como

$$\begin{array}{lll} \underline{Pv_1}^{\delta\,[\underline{v_1}/DQ]} = \mathbb{V} & \underline{Pv_1}^{\delta\,[\underline{v_1}/SP]} = \mathbb{V} & \underline{Pv_1}^{\delta\,[\underline{v_1}/DT]} = \mathbb{F} \\ \underline{Pv_7}^{\delta\,[\underline{v_7}/DQ]} = \mathbb{V} & \underline{Pv_7}^{\delta\,[\underline{v_7}/SP]} = \mathbb{V} & \underline{Pv_7}^{\delta\,[\underline{v_7}/DT]} = \mathbb{F} \end{array}$$

con lo que resulta inmediato determinar que

$$\underline{\forall v_1 Pv_1}^{\delta} = \mathbb{F} \qquad\qquad \underline{\forall v_7 Pv_7}^{\delta} = \mathbb{F}$$

§ 2.37. Práctica (Valores de las fórmulas). Calcular los valores de estas mismas fórmulas, bajo dos de las evaluaciones propuestas en § 2.32.

§ 2.38. Práctica* (Valores de las fórmulas abreviadas). De acuerdo con las abreviaturas introducidas en § 2.23 (p. 53), y basándose en la definición de verdad, verificar paso por paso una de las siguientes cuatro equivalencias:

$$\underline{\beta \vee \gamma}^{\,\varepsilon} = \begin{cases} \mathbb{V} & \text{si } \underline{\beta}^{\,\varepsilon} = \mathbb{V} \text{ o } \underline{\gamma}^{\,\varepsilon} = \mathbb{V} \\ \mathbb{F} & \text{en cualquier otro caso} \end{cases}$$

$$\underline{\beta \wedge \gamma}^{\,\varepsilon} = \begin{cases} \mathbb{V} & \text{si } \underline{\beta}^{\,\varepsilon} = \mathbb{V} \text{ y } \underline{\gamma}^{\,\varepsilon} = \mathbb{V} \\ \mathbb{F} & \text{en cualquier otro caso} \end{cases}$$

$$\underline{\beta \leftrightarrow \gamma}^{\,\varepsilon} = \begin{cases} \mathbb{V} & \text{si } \underline{\beta}^{\,\varepsilon} = \underline{\gamma}^{\,\varepsilon} \\ \mathbb{F} & \text{en caso contrario} \end{cases}$$

$$\underline{\exists x \beta}^{\,\varepsilon} = \begin{cases} \mathbb{V} & \text{si existe algún } a \in U \text{ tal que } \underline{\beta}^{\,\varepsilon\,[x/a]} = \mathbb{V} \\ \mathbb{F} & \text{en caso contrario} \end{cases}$$

§ 2.39. Observación (Problemas 2 y 3). Este es el momento apropiado para plantearse la respuesta a los *Problemas 2* y *3* del listado de *Problemas propuestos* (p. 12).

La consecuencia lógica

§ 2.40. Definición (Verdad lógica). Sea ε una evaluación para un lenguaje de primer orden $\mathscr{L}$:

1. ε *satisface* una fórmula $\underline{\alpha}$ de $\mathscr{L}$ cuando el valor de $\underline{\alpha}$ bajo ε es verdadero, lo cual representamos poniendo "$\varepsilon \vDash \underline{\alpha}$". En caso contrario ponemos "$\varepsilon \nvDash \underline{\alpha}$".

2. Decimos que una fórmula $\underline{\alpha}$ de $\mathscr{L}$ es *"satisfacible"* cuando hay alguna evaluación para $\mathscr{L}$ que *satisface* $\underline{\alpha}$. En caso contrario decimos que es *"insatisfacible"*.

3. Una fórmula $\underline{\alpha}$ de $\mathscr{L}$ es una *lógicamente verdadera* (una *verdad lógica*) cuando es satisfecha por cualquier evaluación para $\mathscr{L}$, lo cual representamos poniendo "$\vDash \underline{\alpha}$". En caso contrario, es decir, en caso de que haya alguna evaluación que no satisfaga $\underline{\alpha}$, ponemos "$\nvDash \underline{\alpha}$".

4. ε *satisface* un conjunto Φ de fórmulas de $\mathscr{L}$ cuando el valor de cualquier fórmula de Φ bajo ε es verdadero, lo cual representamos poniendo "$\varepsilon \vDash \Phi$". En caso contrario ponemos "$\varepsilon \nvDash \Phi$".

5. Un conjunto Φ de fórmulas de $\mathscr{L}$ es *"satisfacible"* cuando hay alguna evaluación para $\mathscr{L}$ que *satisface todas* las fórmulas de Φ. En caso contrario decimos que es *"insatisfacible"*.

§ 2.41. Definición (Consecuencia y equivalencia lógicas).

1. Si Φ es cualquier conjunto de fórmulas de un lenguaje de primer orden $\mathscr{L}$, y $\underline{\alpha}$ una fórmula cualquiera de $\mathscr{L}$, entonces decimos que $\underline{\alpha}$ es una *"consecuencia lógica"* de Φ cuando sucede que cualquier evaluación que satisfaga Φ también satisface $\underline{\alpha}$, lo cual representamos poniendo "$\Phi \vDash \underline{\alpha}$". En caso contrario ponemos "$\Phi \nvDash \underline{\alpha}$".

2. Por su parte, si Φ y Ψ son conjuntos de fórmulas de un lenguaje de primer orden, y resulta que todas y cada una de las fórmulas de Ψ son consecuencias lógicas de Φ, entonces decimos que el conjunto Ψ entero es una *"consecuencia lógica"* de Φ, lo cual representamos poniendo "$\Phi \vDash \Psi$". Y en caso contrario ponemos "$\Phi \nvDash \Psi$".

3. Por último, si $\underline{\beta}$ es también una fórmula de $\mathscr{L}$, entonces decimos que $\underline{\alpha}$ y $\underline{\beta}$ son *"lógicamente equivalentes"* si cualquier evaluación para $\mathscr{L}$ que satisface $\underline{\alpha}$ también satisface $\underline{\beta}$, y viceversa. Es decir: si $\underline{\alpha}$ y $\underline{\beta}$ son satisfechas exactamente por las mismas evaluaciones de $\mathscr{L}$.

§ 2.42. Observación (Consecuencias del conjunto vacío). Como es fácil apreciar, de la definición precedente se sigue de inmediato que una fórmula $\underline{\alpha}$ es consecuencia lógica del conjunto vacío de fórmulas $\varnothing$ si y sólo si es una verdad lógica. Esto es:

$$\varnothing \vDash \underline{\alpha} \qquad \text{si y sólo si} \qquad \vDash \underline{\alpha}$$

§ 2.43. Definición (Término cerrado). Un término es *cerrado* si no contiene ninguna ocurrencia de variables.

§ 2.44. Lema (Valor de los términos cerrados). *Si ε y ε' son cualesquiera evaluaciones para un mismo lenguaje de primer orden $\mathscr{L}$, y están basadas en un mismo modelo, entonces el valor de cualquier término cerrado $\underline{t}$ bajo ε y ε' ha de ser siempre el mismo.*

Prueba. Inmediata por inducción fuerte sobre el grado de $\underline{t}$, distinguiendo las tres formas distintas que puede tomar $\underline{t}$ de acuerdo con la definición de *término.*

En primer lugar, $\underline{t}$ no puede ser una variable, al ser un término cerrado. Y si es una constante, entonces las evaluaciones ε y ε' han de coincidir en el valor que le asignan, al estar basadas en el mismo modelo.

Por último, si $\underline{t}$ tiene la forma

$$\underline{ft_1t_2\dots t_n}$$

donde $\underline{f}$ es un símbolo funcional n-ario de $\mathscr{L}$ y $\underline{t_1}$ $\underline{t_2}$ $\dots\underline{t_n}$ son términos de $\mathscr{L}$, entonces al ser obviamente estos términos cerrados, por hipótesis de inducción tenemos:

$$\underline{t_1}^{\,\varepsilon} = \underline{t_1}^{\,\varepsilon'} \qquad \underline{t_2}^{\,\varepsilon} = \underline{t_2}^{\,\varepsilon'} \qquad \dots \qquad \underline{t_n}^{\,\varepsilon} = \underline{t_n}^{\,\varepsilon'}$$

Como ε y ε' están basadas en el mismo modelo, tenemos también $\underline{f}^{\,\varepsilon} = \underline{f}^{\,\varepsilon'}$, es decir: las dos evaluaciones asignan al símbolo funcional $\underline{f}$ una misma función n-aria.

Y el resultado de aplicar esta función a los valores de $\underline{t_1}$ $\underline{t_2}$ $\dots\underline{t_n}$, será por lo tanto el mismo objeto, con lo cual:

$$\underline{ft_1t_2\dots t_n}^{\,\varepsilon} = \underline{ft_1t_2\dots t_n}^{\,\varepsilon'}$$

§ 2.45. Observación (Valor de los términos cerrados). El lema precedente nos autoriza a asignar directamente a cada término cerrado un valor *bajo el modelo en cuestión*, utilizando para ello el valor que le asigne *cualquier* evaluación basada en ese modelo, y con la seguridad de que dicho valor será invariante con respecto a la evaluación escogida en cada caso para calcularlo.

§ 2.46. Definición (Valores de los términos cerrados). Si $\underline{t}$ es un término cerrado de un lenguaje de primer orden $\mathscr{L}$, entonces el *valor*, $\underline{t}^{\mathscr{M}}$, del término $\underline{t}$ bajo el modelo $\mathscr{M}$, es el valor de $\underline{t}$ en cualquier evaluación basada en $\mathscr{M}$.

§ 2.47. Definición (Ocurrencias libres y ligadas). Una ocurrencia *ligada* de una variable $\underline{x}$ en una fórmula es aquella en que $\underline{x}$ ejerce como variable de cuantificación de un cuantificador universal, o está dentro del alcance de un cuantificador universal que tiene a $\underline{x}$ como variable de cuantificación. Y una ocurrencia *libre* de $\underline{x}$ en una fórmula es aquella en que no suceden ninguna de estas dos cosas.

En general, las *variables libres* de una fórmula son aquellas que tienen alguna ocurrencia libre dentro de esa fórmula. Y las *variables libres* de un conjunto de fórmulas son aquellas que tienen alguna ocurrencia libre dentro de alguna fórmula de ese conjunto.

Por ejemplo, siguiendo con el lenguaje $\mathscr{K}$ de § 2.8, en la fórmula

$$\underline{\forall v_7 P v_7}$$

las dos ocurrencias de la variable $\underline{v_7}$ están ligadas: la primera por actuar como variable de cuantificación, y la segunda por pertenecer al alcance de ese mismo cuantificador. Sin embargo, en la fórmula

$$\underline{\forall v_1 v_2 v_3 (P v_7 \to \neg(v_7 \neq d))}$$

las dos ocurrencias de $\underline{v_7}$ son libres, ya que los cuantificadores en cuyo alcance se encuentran tienen como variables de cuantificación a variables distintas de $\underline{v_7}$. Así como también son libres, lógicamente, las ocurrencias de $\underline{v_7}$ en

$$\underline{P v_7} \qquad \underline{P v_7 \to P d} \qquad \underline{P v_7 \to \neg(v_7 \neq d)}$$

§ 2.48. Definición (Oración). Una *oración* de un lenguaje de primer orden es una fórmula que no posee ninguna variable libre, es decir: que no contiene ninguna ocurrencia libre de una variable.

§ 2.49. Teorema (Valor de las oraciones). *Si ε y ε' son cualesquiera evaluaciones para un mismo lenguaje de primer orden $\mathscr{L}$, y están basadas en un mismo modelo, entonces el valor de cualquier oración $\underline{\alpha}$ bajo ε y ε' ha de ser siempre el mismo.*

Prueba. La prueba de este teorema no es muy intrincada, pero sí algo larga y tediosa, por lo que la omitimos.

§ 2.50. Observación (Valores de las oraciones). El teorema precedente nos autoriza a asignar directamente a cada oración un valor *en el modelo en cuestión*, utilizando para ello el valor que le asigne *cualquier* evaluación basada en ese modelo, y con la seguridad de que dicho valor será invariante con respecto a la evaluación escogida en cada caso para calcularlo.

§ 2.51. Definición (Valor de las oraciones). Si $\underline{\alpha}$ es una oración de un lenguaje de primer orden $\mathscr{L}$, entonces el *valor*, $\underline{\alpha}^{\mathscr{M}}$, de la oración $\underline{\alpha}$ en un modelo $\mathscr{M}$ para $\mathscr{L}$, es el valor de $\underline{\alpha}$ en cualquier evaluación basada en $\mathscr{M}$.

§ 2.52. Práctica (Valores de las oraciones). Proponer dos ejemplos distintos de oraciones del lenguaje $\mathscr{K}$ de § 2.8 (p. 46). A continuación, calcular los valores de cada una en el modelo $\mathscr{S}$ descrito en § 2.27, utilizando para ello cualquier de las dos evaluaciones indicadas en § 2.29.

La sustitución

§ 2.53. Definición (Sustitución en términos). Sean $\underline{r}$ y $\underline{t}$ términos cualesquiera de un lenguaje de primer orden $\mathscr{L}$, y sea $\underline{x}$ cualquier variable de $\mathscr{L}$. El resultado,

$$\underline{(r\,[x/t])}$$

de *sustituir* a $\underline{x}$ por $\underline{t}$ en el término $\underline{r}$, es lo que se obtiene al reemplazar simultáneamente todas las ocurrencias de $\underline{x}$ en $\underline{r}$ por el término $\underline{t}$.

Cuando quiera que un término *sustituido* $\underline{(r\,[x/t])}$ se nombre solo, es decir, no como subcadena de otro término o fórmula, convenimos en que el término $\underline{r\,[x/t]}$ sin más, será por definición igual al término $\underline{(r\,[x/t])}$. Esto nos permite suprimir los paréntesis más exteriores cada vez que mencionamos un término sustituido.

Por ejemplo, volviendo a nuestro consabido lenguaje $\mathscr{K}$ de § 2.8, tenemos, entre otros:

$$\begin{aligned}
\underline{v_1\,[v_1/d]} &= \underline{d}\\
\underline{gv_7\,[v_7/gd]} &= \underline{ggd}\\
\underline{hv_1gv_7\,[v_7/hv_7hdd]} &= \underline{hv_1ghv_7hdd}\\
\underline{gv_7\,[v_1/d]} &= \underline{gv_7}
\end{aligned}$$

Como podemos apreciar, en el último caso la sustitución realizada es vacua, queda sin efecto. Ello se debe a que en el término original, $\underline{gv_7}$, no aparece la variable a ser reemplazada en la sustitución, $\underline{v_1}$. Por ello, el término obtenido al efectuar la sustitución $\underline{[v_1/d]}$ en el término $\underline{gv_7}$, es el propio término $\underline{gv_7}$.

También debemos notar cómo en el caso de la sustitución $\underline{hv_1gv_7\,[v_7/hv_7hdd]}$, es importante que se reemplacen *simultáneamente* todas las ocurrencias de la variable $\underline{v_7}$ en el término original, tal y como reza la definición. Y ello es así, a pesar de que esa variable sólo ocurre una vez en dicho término (!).

La razón es que el término que entra en la sustitución en ese caso, resulta que también contiene ocurrencias de la misma variable de sustitución, $\underline{v_7}$. Y la sustitución indicada, como es lógico, consiste en reemplazar únicamente las ocurrencias de la variable $\underline{v_7}$ que estuviesen contenidas en el término original. Pero no las nuevas ocurrencias que puedan entrar como consecuencia de la propia sustitución efectuada.

§ 2.54. Lema (La sustitución en términos). *Sean:*

- $\underline{r}$ *y* $\underline{t}$ *términos de un lenguaje de primer orden* $\mathscr{L}$
- $\underline{x}$ *una variable de* $\mathscr{L}$
- ε *una evaluación para* $\mathscr{L}$*, y*
- a *el objeto que* ε *asigna a* $\underline{t}$*, es decir:* $a = \underline{t}^{\,\varepsilon}$

Entonces

$$\underline{r\,[x/t]}^{\,\varepsilon} \;=\; \underline{r}^{\,\varepsilon\,[\underline{x}/a]}$$

Prueba. Lo que viene a decir este lema, en otras palabras, es que el valor del término sustituido $\underline{r\,[x/t]}$ bajo ε es el mismo que el valor del término original $\underline{r}$ bajo la evaluación reasignada $\varepsilon\,[\underline{x}/a]$, donde a es aquel elemento del universo que de hecho ε asigna a $\underline{t}$.

La prueba sin embargo es también algo engorrosa, y la omitimos.

Un ejemplo sencillo de la aplicación de este lema, volviendo a nuestra evaluación δ de §2.29 (p. 56) lo tenemos en

$$\underline{v_1\,[v_1/d]}^{\,\delta} \;=\; \underline{d}^{\,\delta} \;=\; DT \;=\; \underline{v_1}^{\,\delta\,[\underline{v_1}/DT,]}$$

§ 2.55. Definición (Sustitución en fórmulas). Sea $\underline{t}$ cualquier término de un lenguaje de primer orden $\mathscr{L}$, α una fórmula, y $\underline{x}$ cualquier variable de $\mathscr{L}$. Definimos el resultado,

$$\underline{(\alpha\,[x/t])}$$

de sustituir a $\underline{x}$ por $\underline{t}$ en la fórmula $\underline{\alpha}$ por inducción fuerte sobre el grado de $\underline{\alpha}$, distinguiendo las cuatro formas distintas que puede tomar $\underline{\alpha}$ de acuerdo con la definición de *fórmula.*

1. Si $\underline{\alpha}$ es una fórmula atómica entonces $\underline{(\alpha\,[x/t])}$ es el resultado de reemplazar simultáneamente todas las ocurrencias de $\underline{x}$ en $\underline{\alpha}$ por el término $\underline{t}$.
2. Si $\underline{\alpha}$ es una fórmula de negación $\underline{\neg\beta}$ donde $\underline{\beta}$ es alguna otra fórmula de $\mathscr{L}$, entonces

   $$\underline{(\neg\beta\,[x/t])}$$

   es sencillamente el resultado de anteponer el símbolo de negación a la fórmula sustituida $\underline{(\beta\,[x/t])}$.
3. Si $\underline{\alpha}$ es una fórmula condicional $\underline{\beta \to \gamma}$, donde $\underline{\beta}$ y $\underline{\gamma}$ son fórmulas de $\mathscr{L}$, entonces

   $$\underline{(\beta \to \gamma\,[x/t])}$$

   es sencillamente el resultado de anteponer el condicional a las fórmulas sustituidas $\underline{(\beta\,[x/t])}$ y $\underline{(\gamma\,[x/t])}$.
4. Y finalmente, si $\underline{\alpha}$ es una fórmula de cuantificación universal

   $$\underline{\forall y \beta}$$

   donde $\underline{y}$ es una variable y $\underline{\beta}$ otra fórmula, ambas de $\mathscr{L}$, entonces distinguimos entre tres posibilidades bien diferenciadas:

(a) Si la variable de cuantificación $\underline{y}$ es idéntica a $\underline{x}$, entonces la sustitución quedará sin efecto:

$$\underline{(\forall x\beta\,[x/t])} = \underline{\forall x\beta}$$

(b) Si la variable de cuantificación $\underline{y}$ es distinta de $\underline{x}$, y además la variable $\underline{y}$ *no ocurre* en el término $\underline{t}$, entonces $\underline{(\forall y\beta\,[x/t])}$ será el resultado de anteponer a la fórmula $\underline{\beta\,[x/t]}$ el cuantificador universal, y la variable $\underline{y}$ como variable de cuantificación:

$$\text{si} \quad \underline{x} \neq \underline{y} \quad \text{e} \quad y \text{ no ocurre en } \underline{t}, \quad \text{entonces} \quad \underline{(\forall y\,\beta[x/t]\,)} = \underline{\forall y\,(\beta\,[x/t])}$$

(c) Por último, si la variable de cuantificación $\underline{y}$ es distinta de $\underline{x}$, pero el término $\underline{t}$ contiene ocurrencias de esta variable $\underline{y}$, entonces tomamos la primera variable de $\mathscr{L}$ que:

- sea distinta de $\underline{x}$ e $\underline{y}$
- no ocurra *en absoluto* ni en $\underline{t}$ ni en $\underline{\beta}$

y, siendo $\underline{z}$ esta variable, efectuamos las sustituciones sucesivas:

$$\underline{(\beta\,[y/z])} \qquad\qquad \underline{(\,(\beta\,[y/z])\,[x/t]\,)}$$

y anteponemos al resultado obtenido el cuantificador existencial con la variable $\underline{z}$ como variable de cuantificación. En definitiva:

$$\underline{(\forall y\,\beta[x/t]\,)} = \underline{\forall z\,(\,(\beta\,[y/z])\,[x/t]\,)}$$

Por último, cuando quiera que una fórmula sustituida $\underline{(\alpha\,[x/t])}$ se nombre sola, es decir, no como subcadena de otra fórmula, convenimos en que la fórmula $\underline{\alpha\,[x/t]}$ sin más, será por definición igual a la fórmula $\underline{(\alpha\,[x/t])}$. Esto nos permite suprimir los paréntesis más exteriores cada vez que mencionamos una fórmula sustituida.

§ 2.56. Definición (Término "libre para"). Sea $\underline{t}$ cualquier término de un lenguaje de primer orden $\mathscr{L}$, y sean $\underline{x}$ cualquier variable y $\underline{\alpha}$ una fórmula, ambas de $\mathscr{L}$. Entonces $\underline{t}$ está *libre para* sustituir a $\underline{x}$ en $\underline{\alpha}$ si $\underline{x}$ no tiene ocurrencias libres en $\underline{\alpha}$ que estén bajo el alcance de un cuantificador cuya variable de cuantificación aparezca en $\underline{t}$.

Obviamente si $\underline{t}$ está *libre para* $\underline{x}$ en $\underline{\alpha}$ entonces la sustitución $\underline{\alpha\,[x/t]}$ se podrá efectuar sin ningún cambio en las variables de cuantificación. Es decir: podemos sustituir las subfórmulas de $\underline{\alpha}$, y en particular las cuantificaciones universales, sin tener que recurrir en ningún caso al procedimiento (c) de la definición de la sustitución en fórmulas.

§ 2.57. Observación (La sustitución en fórmulas). No hay espacio suficiente para discutir aquí por qué la definición de la sustitución en fórmulas ha de ser tan complicada. El caso es que resulta necesario.

Tampoco será preciso, para seguir esta asignatura, ser un gran experto en efectuar sustituciones. En fin, la mejor manera de familiarizarse un poco con ella es practicándola, y a tal efecto, se relacionan a continuación algunos ejemplos, siempre sacados de nuestro infatigable lenguaje $\mathscr{K}$ de § 2.8:

$$\begin{array}{rcl}
\underline{Pv_7\,[v_7/v_1]} & = & \underline{Pv_1} \\
\underline{Pv_7 \to Pv_1\,[v_7/hdv_5]} & = & \underline{Phdv_5 \to Pv_1} \\
\underline{\forall v_7 Pv_7\,[v_7/v_1]} & = & \underline{\forall v_7 Pv_7} \\
\underline{\forall v_1(Pv_7 \to Pv_1)\,[v_7/hdv_5]} & = & \underline{\forall v_1(Phdv_5 \to Pv_1)} \\
\underline{\forall v_1(Pv_7 \to Pv_1)\,[v_7/hdv_1]} & = & \underline{\forall v_2(Phdv_1 \to Pv_2)}
\end{array}$$

Es claro que, en el caso de la fórmula $\underline{\forall v_1(Pv_7 \to Pv_1)}$, el término $\underline{hdv_5}$ está libre en ella para la variable $\underline{v_7}$, pero no así el término $\underline{hdv_1}$, por lo que la sustitución por este último requiere un cambio previo en la variable de cuantificación.

§ 2.58. Práctica (Sustituciones en fórmulas). Siguiendo la línea de los ejemplos anteriores, indicar tres ejemplos alternativos de sustituciones efectuadas en fórmulas, en los que haya que realizar cambios similares a algunos de ellos.

§ 2.59. Teorema (La sustitución en fórmulas). *Sea $\underline{t}$ cualquier término de un lenguaje de primer orden $\mathscr{L}$, y sean $\underline{x}$ cualquier variable y $\underline{\alpha}$ una fórmula, ambas de $\mathscr{L}$. Sea además ε cualquier evaluación para $\mathscr{L}$, y a el objeto que ε asigna a $\underline{t}$, es decir: $a = \underline{t}^{\,\varepsilon}$. Entonces:*

$$\underline{\alpha\,[x/t]}^{\,\varepsilon} = \underline{\alpha}^{\,\varepsilon\,[\underline{x}/a]}$$

Prueba. Lo que viene a decir este teorema, en otras palabras, es que el valor de la fórmula sustituida $\underline{\alpha\,[x/t]}$ bajo una evaluación ε es el mismo que el valor de la fórmula original $\underline{\alpha}$ bajo la evaluación reasignada $\varepsilon\,[\underline{x}/a]$, donde a es el elemento del universo que de hecho ε asigna a $\underline{t}$.

Tampoco probamos aquí este resultado, cuya demostración, esta vez sí es definitivamente larga y compleja. Por poner un ejemplo sencillo del mismo, basta con verificar, tomando esta vez la δ' de § 2.29 (p. 56), que:

$$\underline{Pv_7\,[v_7/v_1]}^{\,\delta'} = \underline{Pv_1}^{\,\delta'} = \mathbb{F} = \underline{Pv_7}^{\,\delta'\,[\underline{v_7}/\tilde{n}]}$$

lo cual es inmediato, dado que $\underline{v_1}^{\,\delta'} = DT$, y DT *no* cumplía la propiedad que correspondía a $\underline{P}$ bajo δ' (cf. § 2.27, p. 55).

El cálculo de predicados

§ 2.60. Definición (Cálculo de predicados). El *cálculo de predicados de primer orden* es un conjunto de reglas que nos permite, dentro de un lenguaje de primer orden, obtener unas fórmulas a partir de otras. Estas reglas se asientan sobre el concepto de *axioma del cálculo de predicados*, la operación de *modus ponens* y el concepto de *deducción*.

La caracterización de cada uno de estos ingredientes conlleva cierta dificultad, por lo que efectuaremos las definiciones correspondientes por separado.

§ 2.61. Definición (Axioma del cálculo de predicados). Un *axioma* del cálculo de predicados de primer orden de un lenguaje $\mathscr{L}$ es una fórmula de $\mathscr{L}$ que pertenece a alguno de los siguientes 11 conjuntos o grupos:

- *Grupo 1.* Lo forman todas las fórmulas de $\mathscr{L}$ que tienen la forma

$$\underline{t = t}$$

  donde $\underline{t}$ es un término cualquiera de $\mathscr{L}$.

- *Grupo 2.* Lo forman todas las fórmulas de $\mathscr{L}$ que tienen la forma

$$\underline{(\,(r_1 = t_1) \wedge (r_2 = t_2) \wedge \ldots \wedge (r_n = t_n)\,) \;\rightarrow\; (\,f r_1 r_2 \ldots r_n = f t_1 t_2 \ldots t_n\,)}$$

  donde $\underline{r_1}, \underline{r_2}, \ldots, \underline{r_n}$ y $\underline{t_1}, \underline{t_2}, \ldots, \underline{t_n}$ son términos cualesquiera de $\mathscr{L}$, y $\underline{f}$ es cualquier símbolo funcional n-ario de $\mathscr{L}$.

- *Grupo 3.* Lo forman todas las fórmulas de $\mathscr{L}$ que tienen la forma

$$\underline{(\,(r_1 = t_1) \wedge (r_2 = t_2) \wedge \ldots \wedge (r_n = t_n)\,) \;\rightarrow\; (\,R r_1 r_2 \ldots r_n \rightarrow R t_1 t_2 \ldots t_n\,)}$$

  donde $\underline{r_1}, \underline{r_2}, \ldots, \underline{r_n}$ y $\underline{t_1}, \underline{t_2}, \ldots, \underline{t_n}$ son términos cualesquiera de $\mathscr{L}$, y $\underline{R}$ es cualquier símbolo relacional n-ario de $\mathscr{L}$.

- *Grupo 4.* Lo forman todas las fórmulas de $\mathscr{L}$ que tienen la forma

$$\underline{\alpha \rightarrow (\beta \rightarrow \alpha)}$$

  donde $\underline{\alpha}$ y $\underline{\beta}$ son fórmulas cualesquiera de $\mathscr{L}$.

- *Grupo 5.* Lo forman todas las fórmulas de $\mathscr{L}$ que tienen la forma

$$\underline{(\,\alpha \rightarrow (\beta \rightarrow \gamma)\,) \;\rightarrow\; (\,(\alpha \rightarrow \beta) \rightarrow (\alpha \rightarrow \gamma)\,)}$$

  donde $\underline{\alpha}$, $\underline{\beta}$ y $\underline{\gamma}$ son fórmulas cualesquiera de $\mathscr{L}$.

- *Grupo 6.* Lo forman todas las fórmulas de $\mathscr{L}$ que tienen la forma

$$\underline{((\neg\alpha \rightarrow \beta) \wedge (\neg\alpha \rightarrow \neg\beta)) \rightarrow \alpha}$$

  donde $\underline{\alpha}$ y $\underline{\beta}$ son fórmulas cualesquiera de $\mathscr{L}$.

- *Grupo 7.* Lo forman todas las fórmulas de $\mathscr{L}$ que tienen la forma

$$\underline{(\,(\alpha \wedge \beta) \;\rightarrow\; \gamma\,) \;\rightarrow\; (\,\alpha \rightarrow (\beta \rightarrow \gamma)\,)}$$

  donde $\underline{\alpha}$, $\underline{\beta}$ y $\underline{\gamma}$ son fórmulas cualesquiera de $\mathscr{L}$.

- *Grupo 8.* Lo forman todas las fórmulas de $\mathscr{L}$ que tienen la forma

$$\underline{(\,\forall x(\alpha \rightarrow \beta)\,) \;\rightarrow\; (\,\forall x\alpha \rightarrow \forall x\beta\,)}$$

  donde $\underline{x}$ es cualquier variable, y $\underline{\alpha}$ y $\underline{\beta}$ son fórmulas cualesquiera de $\mathscr{L}$.

- *Grupo 9.* Lo forman todas las fórmulas de $\mathscr{L}$ que tienen la forma

$$\underline{\alpha \to \forall x \alpha}$$

donde $\underline{\alpha}$ es cualquier fórmula de $\mathscr{L}$, y $\underline{x}$ cualquier variable que no esté libre en $\underline{\alpha}$.

- *Grupo 10.* Lo forman todas las fórmulas de $\mathscr{L}$ que tienen la forma

$$\underline{\forall x \alpha \to (\, \alpha\, [x/t]\,)}$$

donde $\underline{x}$ es cualquier variable, $\underline{\alpha}$ cualquier fórmula, y $\underline{t}$ cualquier término de $\mathscr{L}$ que esté libre para $\underline{x}$ en $\underline{\alpha}$.

- *Grupo 11.* Lo forman todas las fórmulas de $\mathscr{L}$ que tienen la forma

$$\underline{\forall x_1 x_2 \ldots x_n \alpha}$$

donde las variables $\underline{x_1}$, $\underline{x_2}$, ..., $\underline{x_n}$, para algún número natural $n \geqslant 1$, son variables cualesquiera de $\mathscr{L}$, y $\underline{\alpha}$ es un axioma de $\mathscr{L}$ perteneciente a cualquiera de los diez grupos anteriores.

§ 2.62. Definición (Modus ponens). El *modus ponens*, en latín *modo que pone*, es aquella operación que se aplica a cualquier fórmula condicional de un lenguaje de primer orden, y a su fórmula de antecedente, obteniendo como resultado la fórmula de consecuente.

En otras palabras: si $\underline{\alpha}$ y $\underline{\beta}$ son fórmulas cualesquiera de un lenguaje de primer orden, entonces el *modus ponens* se aplica al par de fórmulas

$$\underline{\alpha \to \beta} \qquad\qquad \underline{\alpha}$$

para obtener como resultado la fórmula

$$\underline{\beta}$$

§ 2.63. Definición (Deducción en el cálculo de predicados). Sea Φ un conjunto de fórmulas de un lenguaje de primer orden $\mathscr{L}$, y sea $\underline{\alpha}$ una fórmula cualquiera de $\mathscr{L}$. Una *deducción* de $\underline{\alpha}$ desde Φ en el cálculo de predicados de primer orden de $\mathscr{L}$, es una secuencia finita de fórmulas de $\mathscr{L}$ cuyo último componente es la fórmula $\underline{\alpha}$, y en la cual cada fórmula, incluida la propia $\underline{\alpha}$, ha de ser:

- un axioma del cálculo de predicados de primer orden de $\mathscr{L}$, o bien:
- un miembro de Φ, o bien:
- el resultado de la aplicación del modus ponens a dos fórmulas anteriores en la secuencia

En este contexto, a las fórmulas del conjunto Φ las denominamos *"premisas"* de la deducción, y a la fórmula $\underline{\alpha}$ la denominamos *"conclusión"* de la deducción.

En resumen, una deducción de $\underline{\alpha}$ desde Φ no es más que una secuencia finita de fórmulas, compuesta íntegramente por axiomas, premisas y fórmulas obtenidas de las anteriores por modus ponens, y que tiene como último componente a la conclusión $\underline{\alpha}$.

Por ejemplo, volviendo a nuestro mesiánico lenguaje de § 2.8, y tomando el par de fórmulas de $\mathscr{K}$

$$\Psi = \{\, \underline{v_7 = d}\,,\ \underline{Pv_7}\,\}$$

Entonces, la siguiente secuencia es una deducción de la fórmula $\underline{Pd}$ desde el conjunto Ψ, en el cálculo de predicados de primer orden:

$$\begin{array}{l}
\underline{(v_7 = d) \to (Pv_7 \to Pd)} \\
\underline{v_7 = d} \\
\underline{Pv_7 \to Pd} \\
\underline{Pv_7} \\
\underline{Pd}
\end{array}$$

Como es inmediato comprobar, la primera fórmula de esta deducción es un axioma del cálculo de predicados de $\mathscr{L}$, perteneciente al grupo 3, las fórmulas segunda y cuarta son premisas, y las fórmulas tercera y quinta proceden por modus ponens de parejas de fórmulas anteriores en la deducción.

§ 2.64. Práctica (Deducciones en el cálculo de predicados). Utilizando el mismo lenguaje $\mathscr{K}$, poner otros dos ejemplos de deducciones en el cálculo de predicados, anotando al margen de cada línea su procedencia.

§ 2.65. Definición (Deducibilidad en el cálculo de predicados). Sea α una fórmula cualquiera de un lenguaje de primer orden $\mathscr{L}$.

1. Dado un conjunto Φ de fórmulas de $\mathscr{L}$, decimos que $\underline{\alpha}$ es *"deducible"* de Φ en el cálculo de predicados, si existe alguna deducción de $\underline{\alpha}$ desde Φ en dicho cálculo, lo cual representamos poniendo: "$\Phi \vdash \underline{\alpha}$". Y en caso contrario ponemos "$\Phi \nvdash \underline{\alpha}$".

2. Decimos $\underline{\alpha}$ es un *"teorema"* del cálculo de predicados de primer orden, si existe alguna deducción de $\underline{\alpha}$ desde el conjunto vacío de premisas $\varnothing$, en dicho cálculo, lo cual representamos poniendo: "$\vdash \underline{\alpha}$". Y en caso contrario ponemos sencillamente: "$\nvdash \underline{\alpha}$".

3. Por último, si Φ y Ψ son conjuntos cualesquiera de fórmulas de $\mathscr{L}$, decimos que Φ es *"deducible"* de Ψ en el cálculo de predicados, si son *deducibles* de Φ todas y cada una de las fórmulas del conjunto Ψ, lo cual representamos poniendo: "$\Phi \vdash \Psi$". Y en caso contrario ponemos: "$\Phi \nvdash \Psi$".

§ 2.66. Teorema (Teorema de corte). *Sea $\mathscr{L}$ un lenguaje de primer orden, y sean Δ, Φ y Ψ, conjuntos cualesquiera de fórmulas de $\mathscr{L}$. Entonces, si Φ es deducible de Δ, y a su vez Ψ es deducible de Φ, tenemos necesariamente que Ψ es deducible de Δ. Es decir:*

$$\textit{si} \qquad \Delta \vdash \Phi \qquad \textit{y} \qquad \Phi \vdash \Psi \qquad \textit{entonces} \qquad \Delta \vdash \Psi$$

Prueba. Sea β cualquier fórmula de Ψ. Como sabemos que $\Phi \vdash \Psi$, podemos tomar cualquier deducción Γ de β desde Φ. Como tal deducción, Γ estará compuesta por una secuencia de fórmulas, cada una de las cuales ha de ser o bien un axioma, o bien un miembro de Φ, o bien una fórmula obtenida por modus ponens de dos fórmulas anteriores.

Consideremos entonces al conjunto de todas las fórmulas de Γ que son miembros de Φ. Ahora bien, como $\Delta \vdash \Phi$, todas estas fórmulas son a su vez deducibles de Δ, así que podemos componer una nueva deducción, Γ', en la cual reemplacemos cada aparición de una fórmula de Φ por su deducción completa desde Δ.

Es inmediato comprobar que el resultado, Γ', seguirá siendo una deducción, que seguirá terminando naturalmente con la fórmula $\underline{\beta}$, y que en definitiva, constituirá una deducción de $\underline{\beta}$ desde Δ. Por lo tanto $\underline{\beta}$ es deducible de Δ.

Y con ello hemos demostrado que cualquier fórmula de Ψ es deducible de Δ, con lo que tenemos $\Delta \vdash \Psi$.

§ 2.67. Teorema (Deducciones desde conjuntos infinitos). *Si una fórmula $\underline{\alpha}$ de un lenguaje de primer orden $\mathscr{L}$ es deducible en el cálculo de predicados de $\mathscr{L}$ desde un conjunto infinito de fórmulas Φ, entonces existe un subconjunto finito Φ' de Φ, tal que $\underline{\alpha}$ es deducible de Φ' en ese mismo cálculo. Es decir:*

$$\textit{si } \Phi \vdash \underline{\alpha} \textit{ entonces } \Phi' \vdash \underline{\alpha} \textit{ para algún subconjunto finito } \Phi' \subseteq \Phi$$

Prueba. Sea Γ cualquier deducción de α desde Φ. Como Γ ha de ser una secuencia *finita* de fórmulas, sólo pueden aparecer en ella un conjunto finito de miembros de Φ, digamos Φ'. Es decir, Φ' es el conjunto de todas aquellos miembros de Φ que aparecen en la secuencia Γ.

Pero entonces, la propia secuencia Γ es una deducción de α desde Φ', con lo que es evidente que $\Phi' \vdash \alpha$.

§ 2.68. Definición (Contradicción, inconsistencia). Una *contradicción* de un lenguaje de primer orden $\mathscr{L}$ es una fórmula de la forma

$$\underline{\alpha \wedge \neg\alpha}$$

para cualquier fórmula $\underline{\alpha}$ de $\mathscr{L}$.

Por su parte, decimos que un conjunto Φ de fórmulas de $\mathscr{L}$ es *"inconsistente"* en el cálculo de predicados, cuando existe alguna contradicción que es deducible de Φ en dicho cálculo. En caso contrario decimos que Φ es *"consistente"*.

§ 2.69. Corolario (Conjuntos inconsistentes infinitos). *Si un conjunto infinito de fórmulas es inconsistente, entonces tiene algún subconjunto finito inconsistente.*

Prueba. Inmediata por el *Teorema* § 2.67.

§ 2.70. Teorema (Teorema de corrección del cálculo). *Sea Φ un conjunto de fórmulas de un lenguaje de primer orden $\mathscr{L}$, y $\underline{\alpha}$ una fórmula de $\mathscr{L}$. Si $\underline{\alpha}$ es deducible de Φ en el cálculo de predicados, entonces $\underline{\alpha}$ es una consecuencia lógica de Φ. Es decir:*

$$\textit{si } \Phi \vdash \underline{\alpha} \textit{ entonces } \Phi \vDash \underline{\alpha}$$

Prueba. La prueba de este teorema no es especialmente difícil, pero sí algo larga, y también la omitimos.

§ 2.71. Teorema (Teorema de completitud del cálculo). *Sea Φ un conjunto de fórmulas de un lenguaje de primer orden $\mathscr{L}$, y $\underline{\alpha}$ una fórmula de $\mathscr{L}$. Si $\underline{\alpha}$ es una consecuencia lógica de Φ, entonces $\underline{\alpha}$ es deducible de Φ en el cálculo de predicados. Es decir:*

$$\textit{si } \Phi \vDash \underline{\alpha} \textit{ entonces } \Phi \vdash \underline{\alpha}$$

Prueba. La prueba de este teorema necesitaría un extenso apartado dedicado en exclusiva a ella. Es tema principal y gran protagonista de los cursos de introducción a la lógica, pero en éste no hay espacio para detenernos en ella.

§ 2.72. Observación (Corrección y completitud del cálculo). Los dos teoremas precedentes son de extraordinaria importancia. Juntos, implican que una fórmula es deducible de un conjunto *si y sólo si* es consecuencia lógica suya. En otras palabras: que hay una correspondencia exacta entre la relación de consecuencia lógica y la relación de deducibilidad en el cálculo de predicados.

Evidentemente ello implica, en particular, que una fórmula es deducible del conjunto vacío de premisas $\varnothing$ si y sólo si es consecuencia lógica de $\varnothing$. O dicho en otras palabras: que una fórmula es un teorema del cálculo de predicados si y sólo si es una verdad lógica.

§ 2.73. Práctica* (Consistencia e inconsistencia). Sea $\mathscr{L}$ un lenguaje de primer orden, y Φ un conjunto de fórmulas en ese lenguaje.

(a) Probar que si Φ es satisfacible, entonces es consistente. O lo que es lo mismo: que si Φ es inconsistente, entonces es insatisfacible.

(b) Probar que si Φ es inconsistente, entonces $\Phi \vdash \beta$ para cualquier fórmula β de $\mathscr{L}$.

(c) Probar que el conjunto vacío de fórmulas $\varnothing$ es consistente.

§ 2.74. Teorema (Teorema de satisfacción). *Si Φ es un conjunto consistente de fórmulas de un lenguaje de primer orden $\mathscr{L}$, entonces existe una evaluación que satisface Φ, y cuya cardinalidad es menor o igual que la del propio $\mathscr{L}$.*

Prueba. La prueba de este teorema deriva del tipo de construcción empleada en la prueba del teorema de completitud, y tampoco la vamos a ver aquí.

§ 2.75. Teorema (Compacidad). *Sea Φ un conjunto de fórmulas de un lenguaje de primer orden $\mathscr{L}$. Entonces, si cualquier subconjunto finito de Φ es satisfacible, también lo es el propio Φ.*

Prueba. Si cualquier subconjunto finito de Φ es satisfacible, entonces, aplicando la *Práctica* § 2.73(a), podemos concluir que cualquier subconjunto finito de Φ es consistente. Por lo tanto, dado el *Corolario* § 2.69, el propio Φ ha de ser forzosamente consistente. Y aplicando el teorema anterior, se sigue de inmediato que Φ es también satisfacible.

§ 2.76. Teorema (Teorema de Löwenheim-Skolem). *Si un conjunto de fórmulas Φ de un lenguaje de primer orden $\mathscr{L}$ es satisfacible, entonces existe, en particular, una evaluación que lo satisface, cuya cardinalidad es menor o igual que la de $\mathscr{L}$.*

Prueba. Este teorema debe su nombre al gran lógico noruego Thoralf Skolem, así como al que primero barruntó la prueba del resultado, el alemán Leopold Löwenheim. Una prueba sencilla del mismo, basada en el teorema de satisfacción, queda como ejercicio que a continuación se propone.

§ 2.77. Práctica⋆ (Teorema de Löwenheim-Skolem). Probar el teorema que se acaba de enunciar.

Teorías

§ 2.78. Definición (Teoría para un conjunto de oraciones). Si Φ es un conjunto de oraciones de un lenguaje de primer orden $\mathscr{L}$, llamamos *"teoría para Φ"* (abreviadamente, "*teo* (Φ)") al conjunto de todas las oraciones de $\mathscr{L}$ que son deducibles de Φ en el cálculo de predicados.

§ 2.79. Observación (Teorías para conjuntos de oraciones). Lo primero que tenemos que observar es que, obviamente, $\Phi \vdash teo\,(\Phi)$.

Pero además, por los teoremas de corrección y completitud (cf. § 2.72), resulta evidente que el conjunto de oraciones que son deducibles de Φ en el cálculo de predicados es igual al conjunto de oraciones que son consecuencias lógicas de Φ. Por lo tanto resulta inmediato que *teo* (Φ) es, además, el conjunto de todas las consecuencias lógicas de Φ, es decir, que tenemos también $\Phi \vDash teo\,(\Phi)$.

Ahora veamos un ejemplo. Tomemos dos oraciones $\underline{\beta}$ y $\underline{\beta \to \gamma}$ de un lenguaje de primer orden, y formemos el conjunto

$$\Psi = \{\, \underline{\beta}\,,\ \underline{\beta \to \gamma}\,\}$$

Entonces tenemos que *teo* (Ψ) habrá de incluir a la oración $\underline{\gamma}$, que es obviamente una consecuencia lógica de las otras dos. Pero también, por ejemplo, a $\underline{\neg\neg\beta}$, que es consecuencia lógica de $\underline{\beta}$; a la propia $\underline{\beta}$ sola, que es trivialmente consecuencia lógica suya; así como a la fórmula $\underline{\neg\neg\gamma}$; a la fórmula $\underline{\neg\gamma \to \neg\beta}$; a la fórmula $\underline{\neg\neg\,\neg\neg\beta}$; a la fórmula $\underline{\neg\neg\,\neg\neg\,\neg\neg\beta}$; etc.

El ejemplo anterior hace evidente que cualquier teoría para un conjunto de oraciones habrá de incluir siempre *infinitas* oraciones.

§ 2.80. Observación (Teorías y el teorema de corte). Por otra parte, sea *teo* (Φ) la teoría para un conjunto de oraciones Φ, y sea β una oración deducible de *teo* (Φ). Entonces tenemos simultáneamente

$$\Phi \vdash teo\,(\Phi) \qquad teo\,(\Phi) \vdash \underline{\beta}$$

con lo cual, aplicando el teorema de corte, podemos concluir que $\Phi \vdash \underline{\beta}$.

Ahora bien. Dado $\Phi \vdash \underline{\beta}$, y el hecho de que *teo* (Φ) es el conjunto de oraciones deducibles de Φ, se sigue de inmediato que $\underline{\beta} \in teo\,(\Phi)$. Esto es, que $\underline{\beta}$ es ya de hecho un miembro de *teo* (Φ).

Y con ello hemos demostrado algo muy extraño: que cualquier oración que sea deducible del conjunto *teo* (Φ) es ya un miembro de este conjunto. O dicho de otro modo, *teo* (Φ) incluye a todas las oraciones que son deducibles de él.

Algo enteramente análogo cabe decir, por similares motivos, de las consecuencias lógicas de *teo* (Φ): si una oración es consecuencia lógica de *teo* (Φ), será deducible de él, y por lo tanto será ya un miembro suyo. En otras palabras, que *teo* (Φ) incluye a todas las oraciones que son consecuencias lógicas suyas.

§ 2.81. Definición (Teoría para un modelo). Por su parte, si $\mathscr{M}$ es un modelo para un lenguaje de primer orden $\mathscr{L}$, entonces llamamos *"'teoría para $\mathscr{M}$"* (abreviadamente, "*teo* ($\mathscr{M}$)") al conjunto de todas las oraciones de $\mathscr{L}$ que son verdaderas en $\mathscr{M}$.

§ 2.82. Observación (Teorías para modelos). Dado cualquier modelo $\mathscr{M}$ para un lenguaje de primer orden $\mathscr{L}$, es evidente, claro está, que $\mathscr{M} \vDash teo\,(\mathscr{M})$.

Además, es obvio que la teoría de cualquier modelo será también siempre un conjunto infinito de oraciones.

Y para terminar, debemos observar que con las teorías para modelos pasa algo enteramente similar a lo que ocurría con las teorías para conjuntos de oraciones, esto es: que cualquier oración que sea consecuencia lógica de $teo\,(\mathscr{M})$ ha de estar ya incluida en $teo\,(\mathscr{M})$, y por lo mismo, que cualquier oración que sea deducible de $teo\,(\mathscr{M})$ también ha de estar ya incluida en $teo\,(\mathscr{M})$.

§ 2.83. Observación (Teorías). Como vemos, las teorías para conjuntos de oraciones y las teorías para modelos comparten una curiosa propiedad: son conjuntos de oraciones que contienen ya a todas las oraciones deducibles del conjunto. O dicho de otro modo: son conjuntos de oraciones para los cuales no existe ninguna oración que sea deducible del conjunto, y que no pertenezca ya a él.

Pues bien: ahora definiremos el concepto de *teoría* a secas, como cualquier conjunto de oraciones que satisfaga eso.

§ 2.84. Definición (Teoría). Un conjunto Δ de oraciones de un lenguaje de primer orden $\mathscr{L}$ es una *teoría* de ese lenguaje, cuando no existe ninguna oración de $\mathscr{L}$ que sea deducible de Δ y que no pertenezca a Δ.

Ello significa, naturalmente, que Δ incluye ya de por sí a todas las oraciones deducibles de Δ. Y consiguientemente, Δ incluirá también a todas las oraciones que sean consecuencias lógicas de Δ, por las mismas razones que ocurría en el caso de las teorías para conjuntos de oraciones.

Queda claro entonces, que tanto las teorías para conjuntos de oraciones, como las teorías para modelos, son también *teorías* sin más, en el sentido que se acaba de definir.

§ 2.85. Definición (Axiomas para una teoría). Finalmente, dada una teoría Δ, llamamos *"conjunto de axiomas para"* Δ a cualquier conjunto de oraciones Φ tal que $teo\,(\Phi) = \Delta$, es decir: tal que la teoría Δ sea una teoría para ese conjunto de oraciones.

En definitiva, dados dos conjuntos de oraciones Φ y Δ, siempre que Δ sea una teoría para Φ, a su vez Φ será un conjunto de axiomas para Δ. Y viceversa: siempre que Φ sea un conjunto de axiomas para Δ, tendremos que Δ es precisamente la teoría para Φ.

§ 2.86. Práctica (Axiomas para una teoría). Volviendo al conjunto Ψ de § 2.79, poner un ejemplo de conjunto de axiomas para la teoría $teo\,(\Psi)$ que sea distinto del propio Ψ.

§ 2.87. Observación (Problema 4). Este es el momento apropiado para plantearse la respuesta al *Problema 4* del listado de *Problemas propuestos* (p. 12).

MÓDULO 3

Modelos no estándar

La aritmética de primer orden

§ 3.1. Observación (La aritmética de primer orden). Una vez presentada la lógica de primer orden, vamos a aplicarla a esa teoría matemática tan elemental, que es la *aritmética*. La aritmética es la parte de las matemáticas que estudia la serie de los números naturales: 0, 1, 2, 3, ..., y las cuatro operaciones básicas entre ellos, suma, resta, multiplicación y división. *A priori* parece que tendría que ser muy fácil someter a formalización lógica una teoría tan sencilla, que hasta los niños son capaces de manejar. Pero la respuesta dista mucho de ser la esperada, como vamos a ver enseguida.

Lo primero que vamos a hacer es definir un lenguaje de primer orden que sea adecuado para representar esta teoría matemática. Esto es: un lenguaje en el que se puedan reproducir los cálculos aritméticos habituales, pero que esté construido siguiendo todas las especificaciones propias de los lenguajes formales de primer orden, tal y como las estudiamos en el Módulo anterior. A este lenguaje lo llamaremos "$\mathscr{A}$".

A continuación presentaremos la estructura de los números naturales con el formato de un modelo para el lenguaje $\mathscr{A}$. A dicho modelo lo llamaremos "$\mathscr{N}$", y será para nosotros el *"modelo estándar"* de nuestro lenguaje.

Sin embargo, como $\mathscr{A}$ es un lenguaje puramente formal, tendrá infinidad de modelos posibles, además de $\mathscr{N}$. Para caracterizar al modelo $\mathscr{N}$ por medio del lenguaje $\mathscr{A}$, tendríamos que componer un conjunto de oraciones de ese lenguaje que representaran exhaustivamente las peculiaridades del modelo $\mathscr{N}$, es decir, que representaran exhaustivamente las propiedades de los números naturales. Y el resultado sería una *caracterización formal* de la estructura de los números naturales por medio de un lenguaje de primer orden.

Pues bien: a desentrañar las distintas vicisitudes de esta singular tarea está dedicado el resto de nuestra asignatura.

§ 3.2. Definición (El lenguaje de la aritmética de primer orden). El lenguaje $\mathscr{A}$ de la aritmética de primer orden es aquel que tiene como conjunto de símbolos peculiares:

- una constante, $\underline{0}$, llamada *"constante cero"*
- un símbolo funcional monario, $\underline{s}$, llamado *"símbolo de sucesión"*
- dos símbolos funcionales binarios, $\underline{+}$ y $\underline{\times}$, denominados *"símbolo de adición"* y *"símbolo de multiplicación"* respectivamente

La razón de que $\mathscr{A}$ no posea símbolos funcionales para la resta y para la división, como tampoco para la exponenciación u otras operaciones aritméticas, es que no resulta necesario, porque todas ellas se pueden representar fácilmente a partir de la suma y el producto.

§ 3.3. Observación (El lenguaje $\mathscr{A}$). Dada la anterior definición, es evidente que todos los términos de $\mathscr{A}$ tendrán indefectiblemente una de las siguientes formas:

1. una variable de $\mathscr{A}$ sola
2. la constante $\underline{0}$ sola
3. la forma $\underline{st}$, donde $\underline{t}$ es otro término de $\mathscr{A}$
4. la forma $\underline{+rt}$, donde $\underline{r}$ y $\underline{t}$ son términos de $\mathscr{A}$
5. la forma $\underline{\times rt}$, donde $\underline{r}$ y $\underline{t}$ son términos de $\mathscr{A}$

§ 3.4. Práctica* (Cardinalidad de $\mathscr{A}$). Probar que $\mathscr{A}$ es un lenguaje enumerable.

§ 3.5. Convención (Abreviaturas para los términos de $\mathscr{A}$). Si $\underline{r}$ y $\underline{t}$ son términos cualesquiera de $\mathscr{L}$, entonces:

- A un término del tipo $\underline{+rt}$ lo denotaremos también poniendo "$\underline{(r+t)}$".
- A un término del tipo $\underline{\times rt}$ lo denotaremos también poniendo "$\underline{(r \times t)}$".

§ 3.6. Definición (Numerales). Por inducción débil definimos, para cada $n \in \mathbb{N}$, un término de $\mathscr{A}$, representado por "$\underline{s_n}$", y al que llamaremos el *"n-ésimo numeral"* de $\mathscr{A}$:

- el numeral $\underline{s_0}$ será la misma constante $\underline{0}$.
- si $\underline{s_n}$ es un numeral de $\mathscr{A}$, entonces $\underline{s_{n+1}}$ será el resultado de anteponer el símbolo de sucesión $\underline{s}$ al numeral $\underline{s_n}$, es decir:

$$\underline{s_{n+1}} = \underline{ss_n}$$

De este modo, tenemos:

$$\underline{s_0} = \underline{0} \qquad \underline{s_1} = \underline{s0} \qquad \underline{s_2} = \underline{ss0} \qquad \underline{s_3} = \underline{sss0} \qquad \text{etc.}$$

En otras palabras: cada numeral $\underline{s_n}$ consistirá en una ocurrencia de la constante $\underline{0}$, precedida por n ocurrencias del símbolo funcional $\underline{s}$.

§ 3.7. Definición (El modelo $\mathscr{N}$). De entre la infinita variedad de modelos para $\mathscr{A}$, sin duda el primero que se nos ocurre es el *modelo aritmético*, que denominaremos "$\mathscr{N}$", y cuyos ingredientes son los siguientes:

- Como universo, el modelo $\mathscr{N}$ tiene al conjunto $\mathbb{N}$ de los números naturales
$$\mathbb{N} = \{\, 0, 1, 2, 3, \dots \}$$
- El valor de la constante $\underline{0}$ en $\mathscr{N}$ es el número 0.
- El valor del símbolo funcional $\underline{s}$ en $\mathscr{N}$ es la función de *sucesión* s (cf. p. 39), tal que para todo $n \in \mathbb{N}$,
$$s(n) = n + 1$$
- El valor del símbolo funcional $\underline{+}$ en $\mathscr{N}$ es la conocida función de adición o suma de números naturales $(+)$, que todos hemos aprendido en la escuela.
- El valor del símbolo funcional $\underline{\times}$ en $\mathscr{N}$ es la no menos célebre función producto o multiplicación de naturales ($\times$ ó $\cdot$), que también conocemos todos, o al menos deberíamos conocer.

Del modelo aritmético $\mathscr{N}$ así definido decimos que es el *"modelo estándar"* para el lenguaje de la aritmética de primer orden $\mathscr{A}$, dada la forma tan evidente en que dicho modelo se ajusta a este particular lenguaje formal.

§ 3.8. Lema (Valor de los numerales en $\mathscr{N}$). *Para todo $n \in \mathbb{N}$ tenemos que*
$$\underline{s_n}^{\mathscr{N}} = n$$

Prueba. Inmediata, por inducción débil sobre n. En primer lugar efectuamos la base, para $n = 0$:
$$\begin{aligned} \underline{s_0}^{\mathscr{N}} &= \underline{0}^{\mathscr{N}} && \text{(por definición de } \textit{numeral}\text{)} \\ &= 0 \end{aligned}$$

Y ahora el paso de inducción, tomando un n cualquiera y suponiendo, por hipótesis de inducción, que $\underline{s_n}^{\mathscr{N}} = n$. Entonces:
$$\begin{aligned} \underline{s_{n+1}}^{\mathscr{N}} &= \underline{ss_n}^{\mathscr{N}} && \text{(por definición de } \textit{numeral}\text{)} \\ &= \underline{s}^{\mathscr{N}}\,(\,\underline{s_n}^{\mathscr{N}}\,) \\ &= s\,(\,\underline{s_n}^{\mathscr{N}}\,) && \text{(dado que } \underline{s}^{\mathscr{N}} = s\text{)} \\ &= s(n) && \text{(por hipótesis de inducción)} \\ &= n + 1 \end{aligned}$$

§ 3.9. Observación (Otros modelos posibles para $\mathscr{A}$). Mucho ojo, porque $\mathscr{N}$ no es, ni mucho menos, el único modelo que tiene el lenguaje $\mathscr{A}$. Por ejemplo, el modelo $\mathscr{T}$ que a continuación se describe es otro caso de modelo para este lenguaje:

- Como universo, $\mathscr{T}$ posee aquel conjunto W del que ya hablamos en § 2.27 (p. 55):

$$W = \{ DQ, SP, DT \}$$

- El valor de la constante $\underline{0}$ en $\mathscr{T}$ es el individuo DT, o sea, la amada Dulcinea:

$$\underline{0}^{\mathscr{T}} = DT$$

- El valor del símbolo funcional $\underline{s}$ en $\mathscr{T}$ es aquella función monaria i tal que:

$$i(DQ) = SP \qquad i(SP) = DT \qquad i(DT) = DQ$$

- El valor del símbolo funcional $\underline{+}$ en $\mathscr{T}$ es aquella función binaria j que siempre da DQ como resultado. Es decir, tal que:

$$j(DQ, DQ) = DQ \qquad j(DQ, SP) = DQ \qquad j(SP, DQ) = DQ \qquad \text{etc.}$$

- El valor del símbolo funcional $\underline{\times}$ en $\mathscr{T}$ es la misma función binaria j que se acaba de indicar.

El valor de los numerales, por ejemplo, en este modelo $\mathscr{T}$, resulta ser muy distinto del que tenían en $\mathscr{N}$:

$$\underline{s_0}^{\mathscr{T}} = DT \qquad \underline{s_1}^{\mathscr{T}} = i(DT) = DQ \qquad \underline{s_2}^{\mathscr{T}} = i(DQ) = SP$$

$$\underline{s_3}^{\mathscr{T}} = i(SP) = DT \qquad \underline{s_4}^{\mathscr{T}} = i(DT) = DQ \qquad \ldots$$

Y una oración como $\underline{0+0=0}$, por ejemplo, resulta ser *falsa* en el modelo $\mathscr{T}$, por la sencilla razón de que $\underline{0}^{\mathscr{T}} = DT$, mientras que $\underline{0+0}^{\mathscr{T}} = j(DT, DT) = DQ$.

§ 3.10. Práctica (Otros modelos posibles para $\mathscr{A}$). Describir un nuevo ejemplo de modelo arbitrario para el lenguaje $\mathscr{A}$, cuyo universo tenga exactamente cuatro elementos distintos. Calcular el valor de los primeros diez numerales en ese modelo, e indicar si la oración $\underline{0+0=0}$ resulta en él verdadera o falsa.

§ 3.11. Definición (Los conjuntos Υ_n). Para cada natural $n \geqslant 1$, llamamos "Υ_n" al conjunto de todas aquellas fórmulas de $\mathscr{A}$ que *no* tengan libre ninguna variable *posterior* a $\underline{v_n}$.

Por ejemplo, la fórmula $\underline{v_7 = sv_1}$ pertenecerá a Υ_7, ya que no tiene libre ninguna variable posterior a $\underline{v_7}$. Por lo mismo, también pertenecerá, naturalmente, a Υ_8, a Υ_9, etc.

§ 3.12. Definición (Las teorías Υ_0 y Λ). Además, llamaremos "Υ_0" al conjunto de todas aquellas fórmulas de $\mathscr{A}$ que no tengan libre ni siquiera a la variable $\underline{v_1}$. Es decir: Υ_0 será el conjunto de todas las *oraciones* de $\mathscr{A}$. Así, fórmulas como $\underline{0+0=0}$ o $\underline{\forall v_7(0=v_7)}$, pertenecerán a Υ_0, ya que no tienen libre a ninguna variable de $\mathscr{A}$ en absoluto.

Es obvio entonces que Υ_0 es una teoría, ya que, conteniendo a todas las oraciones de $\mathscr{A}$, no puede existir, en particular, ninguna oración que sea deducible de Υ_0 y que no pertenezca ya a Υ_0.

Por otra parte, vamos a llamar "Λ" a aquella teoría de $\mathscr{A}$ para el conjunto vacío de oraciones $\varnothing$:

$$\Lambda = teo\,(\varnothing)$$

Como las oraciones deducibles del conjunto vacío de premisas son los teoremas del cálculo de predicados, tenemos que Λ es también el conjunto de todas las oraciones de $\mathscr{A}$ que son teoremas del cálculo de predicados. O también, por consiguiente (cf. la observación sobre la corrección y completitud del cálculo, p. 72), el conjunto de todas las oraciones de $\mathscr{A}$ que son verdades lógicas.

§ 3.13. Observación (Las teorías Υ_0 y Λ). Es obvio que Υ_0 es una teoría inconsistente. Es más: se trata de *la única* teoría inconsistente de $\mathscr{A}$, ya que para cualquier otra teoría inconsistente Δ, por la *Práctica* § 2.73(b) (p. 72), *todas* las oraciones de $\mathscr{A}$ serán deducibles de Δ, y al ser una teoría tendrán que pertenecer a Δ. Con lo cual Δ acaba conteniendo todas las oraciones de $\mathscr{A}$, exactamente igual que Υ_0.

Por lo mismo, Υ_0 será *la mayor* teoría de $\mathscr{A}$, en el sentido de que cualquier teoría de $\mathscr{A}$ ha de estar incluida en Υ_0. Para verificar esto basta con observar que las teorías de $\mathscr{A}$ son conjuntos de oraciones, e Υ_0 directamente las contiene todas.

Por el contrario, la teoría Λ es claramente consistente, ya que es la teoría del conjunto vacío de oraciones, y por la *Práctica* § 2.73(c) (p. 72), el conjunto vacío de oraciones de cualquier lenguaje de primer orden es consistente. Y además, es fácil ver que Λ es *la menor* teoría de $\mathscr{A}$, en el sentido en que ha de estar incluida en cualquier otra teoría. En efecto, toda teoría ha de contener a las oraciones deducibles de sí misma, y por tanto, trivialmente, a las oraciones de Λ, que son deducibles del conjunto vacío de premisas.

§ 3.14. Definición (La teoría Ω). Llamaremos "Ω" a aquella teoría de $\mathscr{A}$ para el modelo $\mathscr{N}$, esto es:

$$\Omega = teo\,(\mathscr{N})$$

Ω engloba, por lo tanto, a todas aquellas oraciones de $\mathscr{A}$ que son verdaderas bajo la interpretación $\mathscr{N}$. O dicho de otro modo: Ω contiene exactamente toda la verdad que se puede decir sobre $\mathscr{N}$ en el lenguaje $\mathscr{A}$.

§ 3.15. Definición (Corrección y completitud con respecto a $\mathscr{N}$). Por otra parte, decimos que una teoría Δ de $\mathscr{A}$ es *"correcta con respecto a $\mathscr{N}$"*, cuando todas las oraciones de Δ son verdaderas en $\mathscr{N}$. O lo que viene a ser lo mismo: cuando Δ está incluida en Ω (es decir, $\Delta \subseteq \Omega$).

Un caso trivial de teoría correcta con respecto a $\mathscr{N}$ es Λ. Claro está que todas las oraciones de Λ son verdaderas en el modelo $\mathscr{N}$, ya que son verdades lógicas, y por tanto serán verdaderas en *cualquier* modelo para el lenguaje $\mathscr{A}$. Sin embargo, hay muchas oraciones que son verdaderas en $\mathscr{N}$ y que Λ no contiene. Como la oración $\underline{0+0=0}$, por ejemplo, que es verdadera en el modelo $\mathscr{N}$, pero no es una verdad lógica.

En contraposición, decimos que una teoría Δ de $\mathscr{A}$ es *"completa con respecto a $\mathscr{N}$"*, cuando todas las oraciones verdaderas en $\mathscr{N}$ pertenecen a Δ. O lo que viene a ser lo mismo: cuando Ω está incluida en ella (es decir, $\Omega \subseteq \Delta$).

Por ejemplo, una teoría que es trivialmente completa con respecto a $\mathscr{N}$ es Υ_0. Ésta contiene, sin duda, todas las oraciones verdaderas en $\mathscr{N}$. Pero el problema es que contiene muchas más, incluyendo todas las falsas también.

La única teoría, en definitiva, que encaja exactamente con $\mathscr{N}$, en el sentido en que contiene *"la verdad, toda la verdad y nada más que la verdad"* que se puede expresar en $\mathscr{A}$ sobre el modelo $\mathscr{N}$, es la teoría Ω.

Modelos no estándar de la aritmética

§ 3.16. Convención (Otros modelos para $\mathscr{A}$). A partir de ahora, si introducimos mediante $\mathscr{N}^\star$ un modelo cualquiera para $\mathscr{A}$, convendremos en que:

- $\mathbb{N}^\star$ será el conjunto que actúe como universo de $\mathscr{N}^\star$, sea cual sea ese conjunto.
- $0^\star$ será aquel objeto del universo que $\mathscr{N}^\star$ asigne a la constante $\underline{0}$, sea cual sea ese objeto.
- $s^\star$ será aquella función monaria que $\mathscr{N}^\star$ asigne al símbolo funcional $\underline{s}$, sea cual sea esa función.
- $+^\star$ será aquella función binaria que $\mathscr{N}^\star$ asigne al símbolo funcional $\underline{+}$, sea cual sea.
- $\times^\star$ será aquella función binaria que $\mathscr{N}^\star$ asigne al símbolo funcional $\underline{\times}$, sea cual sea.

Por ejemplo, volviendo a nuestro modelo $\mathscr{T}$ de § 3.9, está claro que en ese caso $\mathbb{N}^\star$ sería el conjunto $\{\, DQ, SP, DT \,\}$, $0^\star$ sería el individuo DT, $s^\star$ sería la función i, y $+^\star$ y $\times^\star$ coincidirían con la función j allí descrita.

§ 3.17. Definición (Isomorfismo). Si $\mathscr{N}^\star$ es un modelo cualquiera para $\mathscr{A}$, entonces decimos que ι es un *"isomorfismo"* entre $\mathscr{N}$ y $\mathscr{N}^\star$ cuando:

1. ι es una correspondencia biunívoca entre los universos respectivos, $\mathbb{N}$ y $\mathbb{N}^\star$.
2. La pareja del 0 bajo la correspondencia ι es precisamente el objeto $0^\star$, es decir:
$$\iota(0) = 0^\star$$
3. Para todo $n \in \mathbb{N}$, la pareja de $s(n)$ bajo la correspondencia ι es precisamente el mismo objeto que resulta de aplicar la función $s^\star$ a la pareja de n, es decir:
$$\iota\,(\,s(n)\,) \;=\; s^\star\,(\,\iota(n)\,)$$

4. Para cualesquiera $n, m \in \mathbb{N}$, la pareja del número que resulta al efectuar la suma $n + m$, es precisamente el mismo objeto que se obtiene al aplicar la función $+^\star$ a las parejas de n y m, es decir:

$$\iota\,(n + m) \,=\, \iota(n) \,+^\star\, \iota(m)$$

5. Y otro tanto con respecto al producto: para cualesquiera $n, m \in \mathbb{N}$, la pareja del número que resulta al efectuar la multiplicación $n \cdot m$, es precisamente el mismo objeto que se obtiene al aplicar la función $\times^\star$ a las parejas de n y m, es decir:

$$\iota\,(n \cdot m) \,=\, \iota(n) \,\times^\star\, \iota(m)$$

Finalmente, si existe algún isomorfismo entre $\mathscr{N}$ y $\mathscr{N}^\star$, entonces decimos que son *"modelos isomorfos"*.

§ 3.18. Observación (Isomorfismos). Lo primero que salta a la vista en la definición de *isomorfismo* es que los universos de los dos modelos implicados, $\mathbb{N}$ y $\mathbb{N}^\star$, han de tener la misma cardinalidad. Ello es inmediato, ya que el isomorfismo ι tiene que ser una correspondencia biunívoca entre ellos. A esta primera condición hay que añadir las otras cuatro, cuyo efecto es el de trasladar el mismo comportamiento de los elementos de $\mathbb{N}$ bajo las funciones s, $+$ y $\cdot$, a sus respectivas parejas de $\mathbb{N}^\star$ bajo las correspondientes funciones $s^\star$, $+^\star$ y $\times^\star$.

Por ejemplo: dado que dos más tres son cinco, es decir, dado que en el modelo $\mathscr{N}$ tenemos $2 + 3 = 5$, ello obliga a que en el modelo $\mathscr{N}^\star$, el resultado de aplicar la función $+^\star$ a las parejas respectivas de 2 y 3 sea precisamente aquel objeto de $\mathscr{N}^\star$ que constituye la pareja del número 5. Y así con las restantes operaciones de $\mathscr{N}$.

En otras palabras: que al margen de la identidad individual de los objetos involucrados, el comportamiento, dentro de cada modelo, de las parejas respectivas, debe ser exactamente el mismo. Ambos modelos deben tener por tanto idéntica *estructura*.

Como dos familias, para entendernos, que constaran del mismo número de miembros, y en las cuales se dieran exactamente las mismas relaciones de parentesco. Por ejemplo, la familia del 1º A, compuesta por un padre, una madre, una hija mayor y dos gemelos, y la familia del 2º B, compuesta también, casualidades de la vida, por un padre, una madre, una hija mayor y dos gemelos. Aun siendo familias de hecho distintas, tendrían una estructura de parentesco exactamente idéntica. Serían, en un sentido análogo al nuestro, "familias isomorfas".

§ 3.19. Definición (Modelos no estándar para $\mathscr{A}$). La diferencia entre dos modelos isomorfos es tan pequeña que podemos despreciarla, virtualmente a todos los efectos prácticos. Es por ello que consideramos *modelo estándar* para el lenguaje $\mathscr{A}$, tanto al propio modelo $\mathscr{N}$ como, por extensión, a cualquier modelo isomorfo a él. Y por el contrario, llamamos *"modelo no estándar"* del lenguaje $\mathscr{A}$ a cualquier modelo para este lenguaje que no sea isomorfo al modelo $\mathscr{N}$.

Un ejemplo inmediato de modelo no estándar lo tenemos, no cabe duda, en el modelo $\mathscr{T}$ de § 3.9. Su cardinalidad es ya de entrada distinta de la de $\mathbb{N}$, con lo que está claro que resulta imposible que haya ningún isomorfismo entre ambos.

§ 3.20. Definición (Inmersión). El concepto de *inmersión* del modelo $\mathscr{N}$ en otro $\mathscr{N}^\star$ es un poco más amplio que el de isomorfismo. En realidad la definición es casi enteramente igual, pero cambiando la condición de que la correspondencia sea biunívoca entre los universos respectivos, $\mathbb{N}$ y $\mathbb{N}^\star$, por la formulación más laxa, de que sea biunívoca entre $\mathbb{N}$ y *"algún subconjunto"* de $\mathbb{N}^\star$.

En efecto, si $\mathscr{N}^\star$ es un modelo cualquiera para $\mathscr{A}$, entonces ι es una *inmersión de $\mathscr{N}$ en $\mathscr{N}^\star$* cuando:

1. ι es una correspondencia biunívoca entre $\mathbb{N}$ y algún subconjunto de $\mathbb{N}^\star$.

2. $\iota(0) = 0^\star$.

3. Para todo $n \in \mathbb{N}$, tenemos que $\iota\,(\,s(n)\,) \;=\; s^\star\,(\,\iota(n)\,)$.

4. Para cualesquiera $n, m \in \mathbb{N}$, tenemos que $\iota\,(n+m) \;=\; \iota(n) \;+^\star\; \iota(m)$.

5. Y para cualesquiera $n, m \in \mathbb{N}$, tenemos que $\iota\,(n \cdot m) \;=\; \iota(n) \;\times^\star\; \iota(m)$.

Si, en particular, ι es una correspondencia entre $\mathbb{N}$ y un subconjunto *propio* de $\mathbb{N}^\star$, entonces se trata de una *inmersión propia de $\mathbb{N}$ a $\mathbb{N}^\star$*. Y por el contrario, si ι es una correspondencia entre $\mathbb{N}$ y todo $\mathbb{N}^\star$, entonces la inmersión en cuestión es un isomorfismo sin más.

§ 3.21. Observación (Inmersiones propias). La diferencia está, para entendernos, en que una inmersión propia viene a ser algo así como un isomorfismo entre $\mathbb{N}$ y un subconjunto propio de $\mathbb{N}^\star$. Es decir: de tal forma que el modelo $\mathscr{N}$ encuentre un fiel reflejo no en todo $\mathscr{N}^\star$, sino en una fracción parcial suya. No es que los dos modelos compartan la misma estructura, en este caso, sino que la estructura de uno es idéntica a la de una fracción parcial del otro.

Volviendo a nuestra metáfora anterior sobre las familias, es como si la familia del 1º A, compuesta por dos padres, una hija mayor y dos gemelos, fuese comparada con otra más amplia, digamos la del 3º C, formada en este caso por los padres, dos abuelos, tres hijas mayores, dos gemelos, un bebé y un perro. En este caso, la familia del 1º comparte estructura de parentesco con una fracción parcial de la del 3º, pero no con ella entera.

§ 3.22. Lema (Términos cerrados e inmersiones). *Sea $\mathscr{N}^\star$ un modelo cualquiera de $\mathscr{A}$, e ι una inmersión de $\mathscr{N}$ en $\mathscr{N}^\star$. Entonces, para cualquier término cerrado $\underline{t}$ de $\mathscr{A}$, tenemos*

$$\iota\,(\,\underline{t}^{\,\mathscr{N}}\,) \;=\; \underline{t}^{\,\mathscr{N}^\star}$$

Prueba. Lo que se constata en este lema es que si hay una inmersión de $\mathscr{N}$ en $\mathscr{N}^\star$, entonces el valor que reciba cualquier término cerrado $\underline{t}$ en el modelo $\mathscr{N}^\star$, será justamente la pareja que dicha inmersión asigna a su valor en $\mathscr{N}$.

La prueba es sencilla, por inducción fuerte sobre el grado de $\underline{t}$. Para ello tratamos separadamente las distintas formas que puede tomar $\underline{t}$ de acuerdo con la *Observación* § 3.3, salvando el caso de que sea una variable sola, lo cual es imposible al ser $\underline{t}$ un término cerrado.

1. En primer lugar, supongamos que $\underline{t}$ sea la constante $\underline{0}$. En tal caso:
$$\begin{aligned} \iota\,(\underline{0}^{\mathscr{N}}) &= \iota\,(0) && \text{(dado que } \underline{0}^{\mathscr{N}} = 0) \\ &= 0^{\star} && \text{(por definición de } \textit{inmersión}) \\ &= \underline{0}^{\mathscr{N}^{\star}} && \text{(por la } \textit{Convención } \S\,3.16) \end{aligned}$$

2. A continuación, supongamos que $\underline{t}$ sea de la forma $\underline{sr}$, para algún otro término cerrado $\underline{r}$ de $\mathscr{A}$. Entonces:
$$\begin{aligned} \iota\,(\underline{sr}^{\mathscr{N}}) &= \iota\,(\,s\,(\underline{r}^{\mathscr{N}})\,) && \text{(dado que } \underline{s}^{\mathscr{N}} = s) \\ &= s^{\star}\,(\,\iota\,(\underline{r}^{\mathscr{N}})\,) && \text{(por definición de } \textit{inmersión}) \\ &= s^{\star}\,(\underline{r}^{\mathscr{N}^{\star}}) && \text{(por hipótesis de inducción)} \\ &= \underline{sr}^{\mathscr{N}^{\star}} && \text{(por la } \textit{Convención } \S\,3.16) \end{aligned}$$

3. A continuación, supongamos que $\underline{t}$ sea de la forma $\underline{r_1 + r_2}$, donde $\underline{r_1}$ y $\underline{r_2}$ son términos cerrados de $\mathscr{A}$. Entonces:
$$\begin{aligned} \iota\,(\underline{r_1 + r_2}^{\mathscr{N}}) &= \iota\,(\underline{r_1}^{\mathscr{N}} + \underline{r_2}^{\mathscr{N}}) && \text{(dado que } \underline{+}^{\mathscr{N}} \text{ es la función } +) \\ &= \iota\,(\underline{r_1}^{\mathscr{N}}) +^{\star} \iota\,(\underline{r_2}^{\mathscr{N}}) && \text{(por definición de } \textit{inmersión}) \\ &= \underline{r_1}^{\mathscr{N}^{\star}} +^{\star} \underline{r_2}^{\mathscr{N}^{\star}} && \text{(por hipótesis de inducción)} \\ &= \underline{r_1 + r_2}^{\mathscr{N}^{\star}} && \text{(por la } \textit{Convención } \S\,3.16) \end{aligned}$$

4. Y por último, en el caso en que $\underline{t}$ sea de la forma $\underline{r_1 \times r_2}$, se procede de modo enteramente similar al caso precedente.

§ 3.23. Definición (Conjunto de Skolem). En el presente Manual llamaremos *"conjunto de Skolem"* al conjunto Σ de fórmulas que se obtiene al añadir a la teoría Ω todas las fórmulas de la forma
$$\underline{v_1 \neq s_n}$$
para cada numeral $\underline{s_n}$ de $\mathscr{A}$. Es decir:
$$\Sigma = \Omega \cup \{\, \underline{v_1 \neq s_0}\,,\ \underline{v_1 \neq s_1}\,,\ \underline{v_1 \neq s_2}\,,\ \underline{v_1 \neq s_3}\,,\ \ldots \}$$

§ 3.24. Teorema (Subconjuntos finitos de Σ). *Cualquier subconjunto finito de Σ es satisfacible.*

Prueba. Sea Σ' cualquier subconjunto finito de Σ. Siendo finito, Σ' no podrá contener todas las fórmulas de la secuencia
$$\underline{v_1 \neq s_0}\,,\ \underline{v_1 \neq s_1}\,,\ \underline{v_1 \neq s_2}\,,\ \ldots$$

sino sólo una parte finita de ella.

Suponiendo que de hecho no contenga ninguna, entonces Σ' será obviamente satisfacible, al constituir una mera fracción de Ω.

Suponiendo que Σ' sí contenga algunas fórmulas de la mencionada secuencia, tomemos el *mayor* número natural n tal que la correspondiente fórmula $\underline{v_1 \neq s_n}$ pertenece a Σ'.

Y sea a continuación ε cualquier evaluación basada en $\mathscr{N}$ tal que

$$\underline{v_1}^{\,\varepsilon} = n+1$$

Entonces razonamos como sigue. Por el *Lema* § 3.8 (p. 77) sabemos que $\underline{s_n}^{\,\mathscr{N}} = n$, y como ε está basada en $\mathscr{N}$, también tendremos naturalmente

$$\underline{s_n}^{\,\varepsilon} = n$$

Ello implica que $\underline{v_1}^{\,\varepsilon} \neq \underline{s_n}^{\,\varepsilon}$, y por lo tanto la fórmula $\underline{v_1 \neq s_n}$ resultará *verdadera* en la evaluación ε.

Es inmediato comprobar que a las demás fórmulas de la mencionada secuencia que pertenezcan a Σ', les pasará exactamente lo mismo: como han de ser todas anteriores a $\underline{v_1 \neq s_n}$, el valor de $\underline{v_1}^{\,\varepsilon}$, que es $n+1$, será siempre mayor que el valor del numeral correspondiente, y por lo tanto la respectiva fórmula de desigualdad resultará verdadera bajo ε.

Por último, en cuanto a las fórmulas de Ω que puedan pertenecer a Σ', al ser todas verdaderas en el modelo $\mathscr{N}$, y estar ε basada en dicho modelo, resultarán también evidentemente verdaderas en la evaluación ε.

Y con ello hemos demostrado que *todas* las fórmulas del conjunto Σ' resultan satisfechas por la evaluación ε, con lo cual queda a su vez probado que Σ' es satisfacible.

§ 3.25. Corolario (Satisfacibilidad de Σ). *El conjunto de Skolem Σ es satisfacible.*

Prueba. Inmediata por el teorema precedente, aplicando a su vez el teorema de compacidad: dado que cualquier subconjunto finito de Σ es satisfacible, también ha de serlo el propio Σ.

§ 3.26. Teorema (Modelos no estándar de la aritmética). *Existen modelos no estándar de Ω con universo enumerable.*

Prueba. Dado que el conjunto Σ de fórmulas es satisfacible, y que el lenguaje $\mathscr{A}$ al que pertenece es enumerable (*Práctica* § 3.4, p. 76), podemos aplicar el teorema de Löwenheim-Skolem: existe, en particular, una evaluación que satisface Σ, y cuyo universo tiene una cardinalidad menor o igual que $\mathscr{A}$, es decir, finita o enumerable.

Sea entonces $\varepsilon^\star$ cualquiera de esas evaluaciones, y sea a su vez $\mathscr{N}^\star$ el modelo de $\mathscr{A}$ en el cual $\varepsilon^\star$ está basada.

Lo primero que detectamos es que como la evaluación $\varepsilon^\star$ satisface Σ, y Σ contiene todas las oraciones de Ω, la evaluación $\varepsilon^\star$ hará verdaderas también a todas las oraciones de Ω. Y por lo tanto todas las oraciones de Ω serán verdaderas en la interpretación $\mathscr{N}^\star$, o lo que es lo mismo: $\mathscr{N}^\star$ será un modelo de Ω.

A continuación, es fácil ver que el universo de $\mathscr{N}^\star$ no puede ser finito, y que por lo tanto ha de ser forzosamente enumerable. En efecto, $\mathscr{N}^\star$ hace verdaderas, como hemos visto, a todas las oraciones de Ω. Y por lo tanto, en particular, tendrá que hacer verdaderas a todas las oraciones de la forma

$$\underline{s_m \neq s_n}$$

para cualesquiera numerales distintos, $\underline{s_m}$ y $\underline{s_n}$. Y para que todas esas oraciones sean verdaderas en $\mathscr{N}^\star$, dicho modelo tendrá que asignar valores distintos a todos y cada uno de los numerales de $\mathscr{A}$, que como son infinitos, sólo es posible si el universo de $\mathscr{N}^\star$ es infinito también.

En definitiva, ya sabemos que $\mathscr{N}^\star$ es un modelo de Ω, y que tiene un universo enumerable. Sólo nos queda por probar que $\mathscr{N}^\star$ es no estándar, es decir, no isomorfo al modelo $\mathscr{N}$.

Para verificar esto, basta con tomar cualquier inmersión ι de $\mathscr{N}$ en $\mathscr{N}^\star$, y demostrar que ι no puede ser un isomorfismo, esto es, que ι no puede ser una correspondencia biunívoca entre los universos respectivos.

Ello no es difícil. En efecto, como la evaluación $\varepsilon^\star$ satisface Σ, hará verdaderas a todas las fórmulas de la forma

$$\underline{v_1 \neq s_n}$$

para cada numeral $\underline{s_n}$. Ello implica que el valor de $\underline{v_1}^{\varepsilon^\star}$ sea distinto de $\underline{s_n}^{\varepsilon^\star}$ para cualquier numeral $\underline{s_n}$ de $\mathscr{A}$. O dicho de otro modo:

$$\underline{v_1}^{\varepsilon^\star} \neq \underline{s_n}^{\varepsilon^\star} \qquad \text{para todo } n \in \mathbb{N}$$

Ahora bien:

$$\begin{aligned} \underline{s_n}^{\varepsilon^\star} &= \underline{s_n}^{\mathscr{N}^\star} && \text{(al ser } \underline{s_n} \text{ un término cerrado)} \\ &= \iota\,(\underline{s_n}^{\mathscr{N}}) && \text{(por el } \textit{Lema}\ \S\,3.22\text{, p. 82)} \\ &= \iota\,(n) && \text{(por el } \textit{Lema}\ \S\,3.8\text{, p. 77)} \end{aligned}$$

Lo cual demuestra que

$$\underline{v_1}^{\varepsilon^\star} \neq \iota\,(n) \qquad \text{para todo } n \in \mathbb{N}$$

Por consiguiente, el valor que recibe la variable $\underline{v_1}$ bajo $\varepsilon^\star$, que a la fuerza ha de ser un elemento del universo $\mathbb{N}^\star$, no es pareja de ningún elemento de $\mathbb{N}$ bajo la correspondencia ι. Y ello demuestra que en $\mathbb{N}^\star$ hay al menos un elemento sin emparejar. Por lo tanto ι no puede ser una correspondencia biunívoca.

§ 3.27. Observación (Modelos no estándar de la aritmética). Fue Skolem, en 1934, quien definió el conjunto de oraciones que aquí hemos llamado "Σ", y obtuvo el resultado precedente, que es de singular importancia. Como vimos en su momento, Ω recogía exactamente *toda la verdad* que se puede expresar sobre el modelo aritmético $\mathscr{N}$ en nuestro lenguaje de primer orden $\mathscr{A}$. Sin embargo, ahora comprobamos que Ω tiene también otros modelos no isomorfos con $\mathscr{N}$, es decir, otros modelos estructuralmente distintos a $\mathscr{N}$.

Ello sólo puede significar que hay aspectos del modelo $\mathscr{N}$ que no son expresables en el lenguaje $\mathscr{A}$. Y como $\mathscr{A}$ es sin duda el lenguaje de primer orden que mejor encaja con este modelo, podemos concluir que se trata de una limitación inherente a este tipo de lenguajes, es decir: que la lógica de primer orden resulta incapaz de proporcionar una caracterización formal del modelo aritmético $\mathscr{N}$, que lo distinga efectivamente de cualquier otro modelo no isomorfo a él. Ni siquiera, como hemos visto, dentro de los modelos que tienen la misma cardinalidad que $\mathscr{N}$.

Este hecho pone de relieve una limitación de la lógica de primer orden, que es sin duda sorprendente, sobre todo por tratarse de un modelo matemático tan elemental y básico.

Por otra parte, este hecho también nos debe hacer preguntarnos por la naturaleza de dicho modelo, que creemos conocer tan bien, pero al cual, como vemos, no podemos referirnos unívocamente de una manera *formal*, al menos dentro de la lógica de primer orden.

Se trata de un resultado, por lo tanto, que tiene implicaciones no sólo para la filosofía de la lógica, sino también para la filosofía de la matemática, y en especial para la llamada *"escuela formalista"*, que lideró a principios del siglo XIX el gran matemático alemán David Hilbert. La pretensión de reducir las estructuras matemáticas a meros juegos de signos, que es, muy simplificada, una de las aspiraciones de la citada escuela, encuentra aquí un claro obstáculo. En efecto, ni siquiera en el caso de la elemental estructura aritmética, compuesta por los números naturales y las operaciones aritméticas básicas, somos capaces de proporcionar una caracterización formal que la identifique unívocamente, al menos mediante la lógica de primer orden.

Y para terminar pensemos lo siguiente: si nuestra convicción de que la estructura abstracta $\mathscr{N}$ tiene una entidad propia e independiente, distinta a la de cualquier otra estructura matemática, no puede reposar en las propiedades formales de la teoría que la representa, entonces ¿en qué reposa? En otras palabras: ¿en qué nos basamos para creer que efectivamente esa estructura abstracta e infinita existe de un modo diferenciado y por sí misma?

§ 3.28. Observación (La aritmética de segundo orden). No vamos a entrar a exponer aquí cómo la limitación señalada se resuelve, aparentemente, en la llamada *"lógica de segundo orden"*, que es una extensión de gran alcance de la teoría principal. Baste decir que dicho sistema de lógica constituye una teoría bastante más compleja, que excede con mucho a la teoría del razonamiento deductivo correcto, al menos según la mayoría de los autores, y para la cual no hay ni puede haber un cálculo deductivo completo análogo a nuestro cálculo de predicados, lo cual la deja en cierta manera en el aire. Y que además, lo que consigue caracterizar no es tampoco directamente nuestro modelo elemental $\mathscr{N}$, sino una versión a su vez más elaborada y compleja de este modelo, la llamada *"aritmética de segundo orden"*, de la cual el modelo $\mathscr{N}$ es un pequeño fragmento.

§ 3.29. Práctica⋆ (Inmersiones e isomorfismos). Sea $\mathscr{N}^\star$ una interpretación cualquiera del lenguaje de la aritmética de primer orden $\mathscr{A}$, y sea ι una inmersión del modelo estándar $\mathscr{N}$ en $\mathscr{N}^\star$. Para cada evaluación ε basada en $\mathscr{N}$, definimos una evaluación $\varepsilon^\star$ basada en $\mathscr{N}^\star$, estipulando que para cualquier variable $\underline{x}$ de $\mathscr{A}$,

$$\underline{x}^{\varepsilon^\star} = \iota(\underline{x}^{\varepsilon})$$

(a) Probar que para cada variable $\underline{y}$ de $\mathscr{A}$ y número natural n, las evaluaciones $\varepsilon\,[\underline{y}/n]^\star$ y $\varepsilon^\star\,[\underline{y}/\iota(n)]$ son idénticas. (Verificar que están basadas en el mismo modelo, y que los valores que asignan, tanto a $\underline{y}$ como a las variables distintas de $\underline{y}$, siempre coinciden.)

(b) Probar que para cualquier término $\underline{t}$ de $\mathscr{A}$, tenemos $\underline{t}^{\,\varepsilon^\star} = \iota\,(\,\underline{t}^{\,\varepsilon}\,)$. (Proceder por inducción fuerte sobre *grado* ($\underline{t}$).)

(c) Probar que si ι es, en particular, un isomorfismo entre $\mathscr{N}$ y $\mathscr{N}^\star$, entonces para cualquier fórmula $\underline{\alpha}$ de $\mathscr{A}$, tenemos igualmente $\underline{\alpha}^{\,\varepsilon^\star} = \underline{\alpha}^{\,\varepsilon}$. (Proceder por inducción fuerte sobre *grado* ($\underline{\alpha}$).)

§ 3.30. Observación (Problemas 5 y 6). Este es el momento apropiado para plantearse la respuesta a los *Problemas 5* y *6* del listado de *Problemas propuestos* (p. 12).

MÓDULO 4

Nociones de teoría de la recursión

Relaciones y funciones recursivas

§ 4.1. Observación (La teoría de la recursión). Hay principalmente tres formas de presentar los conceptos centrales de la teoría de la recursión, o teoría de la computabilidad. Cada una de estas tres aproximaciones está basada en una noción distinta, aunque las tres nociones son esencialmente equivalentes, e igual de básicas y fundamentales las tres.

La más inmediata de estas tres aproximaciones se apoya en la noción de *instrucciones mecánicas*, o dicho más exactamente, en la noción de *"instrucciones de manipulación simbólica que se pueden ejecutar de un modo rutinario"*. En 1936 el lógico inglés Alan Turing concibió una ingeniosa manera de precisar completamente esta noción, describiendo una máquina ideal de cálculo, con memoria ilimitada, a la cual se van suministrando instrucciones precisas para que efectúe determinadas operaciones.

A dichas máquinas se las llamó *"máquinas de Turing"*, aunque después se han diseñado muchas otras versiones con nombres distintos. Se podría decir que, en algún sentido, estas máquinas fueron precursoras de nuestros actuales ordenadores, aunque el interés en ellas era puramente abstracto, no para efectuar computaciones reales.

Si suprimimos la descripción detallada de la máquina ideal y la forma como trabaja, y nos quedamos con la noción de que hay instrucciones susceptibles de ser ejecutadas de un modo puramente rutinario, podemos reformular los conceptos centrales de la teoría de la recursión en un formato intuitivo y sumamente facilitado. Y eso será justamente lo que haremos nosotros aquí, a lo largo del presente Módulo.

Por otra parte, también en 1936, aunque de forma independiente, el lógico estadounidense Alonzo Church ideó otra manera bien distinta de precisar los mismos conceptos, que se basaba en la utilización de un *cálculo deductivo*, específicamente diseñado para representarlos. Nosotros no estudiaremos aquí ese cálculo, llamado *"cálculo lambda"*, pero sí veremos más adelante un ejemplo de teoría formal, en este sentido equivalente a un cálculo, en la que resultan representables todas las relaciones recursivas (cf. § 6.14, p. 114).

Y por último, la tercera de las grandes aproximaciones a la teoría de la recursión, y la propuesta que en realidad fue la primera en orden cronológico, es la elaborada por el gran lógico de origen austriaco Kurt Gödel en 1934. La definición de Gödel estuvo inspirada por una sugerencia del matemático francés Jacques Herbrand, por lo que se conoce como *"definición de Gödel–Herbrand"*. Esta definición se basa precisamente en una utilización del *principio de inducción matemática* en el que tanto nos venimos apoyando, combinado con otras operaciones básicas mediante unos determinados patrones.

Cuando Church propuso su definición de recursividad, conjeturó que captaba perfectamente la idea intuitiva que quería caracterizar, es decir: que no podía existir una relación que razonablemente pudiera argumentarse que fuera recursiva por otros criterios, y que no encajase con su definición. A dicha conjetura se la conoce desde entonces como *"tesis de Church"*. Poco después se demostró que en realidad las tres definiciones, la de Church, la de Turing, y la de Gödel–Herbrand, eran equivalentes, así como otras variantes que se han propuesto después. Es por ello que la tesis de Church se aplica hoy en día a cualquiera de estas definiciones, gozando de una aceptación prácticamente universal.

§ 4.2. Convención (Relaciones y funciones numéricas). A partir de este momento, y ya en todo lo que resta del presente Manual de Curso, siempre que hablemos de "propiedades", "relaciones" o "funciones" sin mayor especificación, se entenderá que se trata de propiedades, relaciones y funciones *numéricas*, es decir: de propiedades y relaciones relativas a los números naturales, y de funciones cuyos argumentos y valores sean también números naturales.

En efecto, las nociones básicas de la teoría de la recursión se vienen a referir precisamente a las relaciones y funciones numéricas, y ello es así por la sencilla razón de que viene a ser mucho más ventajoso hacerlo de este modo, para poder desarrollar adecuadamente el estudio de dichas nociones.

Resulta muy fácil, sin embargo, extender después a cualquier otro tipo de manipulación simbólica, los resultados obtenidos por la teoría de la recursión con respecto a los números naturales. Y ello se lleva a cabo mediante algún sistema de *codificación*, del alfabeto de símbolos que lenguaje que sea, en números naturales concretos. Como por ejemplo, el método de codificación que describiremos nosotros más adelante, a partir de § 5.15 (p. 100), para aplicar diversos resultados de la teoría de la recursión, al análisis de nuestro lenguaje formal $\mathscr{A}$, y de las teorías formuladas en dicho lenguaje.

§ 4.3. Convención (Notación del vector). Para cada número natural $n \geqslant 1$, denotamos como

$$\text{“}\mathbb{N}^{\,n}\text{”}$$

(leído *"$\mathbb{N}$ súper n"*), al conjunto de todas las secuencias n-arias de números naturales, es decir: al conjunto de todas las secuencias de números naturales que constan de n componentes.

Así, $\mathbb{N}^{\,2}$, por ejemplo, es el conjunto de todas las secuencias binarias de números naturales. $\mathbb{N}^{\,3}$ es el conjunto de todas las secuencias ternarias, etc. Por otra parte, naturalmente $\mathbb{N}^{\,1} = \mathbb{N}$.

Dicho esto, convenimos en que si $\vec{a}$ (leído *"a vector"*) es una secuencia n-aria de números naturales, es decir, si $\vec{a} \in \mathbb{N}^{\,n}$, entonces denotaremos automáticamente a los componentes de $\vec{a}$ como "a_1", "a_2", ..., "a_n", por ese orden.

Por ejemplo, si $\vec{a} \in \mathbb{N}^9$ es la secuencia correspondiente al número de teléfono de mi despacho, entonces tenemos inmediatamente

$$a_1 = 9 \qquad a_2 = 6 \qquad a_3 = 8 \qquad a_4 = 3 \qquad a_5 = 6 \qquad a_6 = 7 \qquad a_7 = 7 \qquad a_8 = 5 \qquad a_9 = 3$$

§ 4.4. Definición (Relación recursiva). Sea R una relación n-aria. Un *procedimiento de decisión* para R es un conjunto finito de instrucciones precisas que, aplicadas a cualquier secuencia $\vec{a}$ de números naturales $\vec{a} \in \mathbb{N}^n$, nos permite decidir, de forma mecánica y en un número finito de pasos, si $R(\vec{a})$ se da o no.

Y decimos que una relación es *"recursiva"* sencillamente cuando existe un procedimiento de decisión para ella.

Por ejemplo, la relación binaria *ser mayor que* es claramente recursiva. Un procedimiento de decisión para verificar si un número natural a_1 es o no mayor que otro a_2, sería, groso modo, el siguiente: primero se toman los números en su representación decimal habitual, y se verifica cuál de los dos tiene más cifras. En caso de que coincidan, se comprueba cuál tiene mayor el dígito más a la izquierda, siguiendo para esta comparación de los dígitos la ordenación básica obvia, del "0" al "9". En caso de que coincidan también, se comprueba el siguiente dígito hacia la izquierda, y si también coinciden el siguiente, etc.

§ 4.5. Definición (Relación recursivamente enumerable). Sea R una relación n-aria. Un *procedimiento de enumeración* para R es un conjunto finito de instrucciones precisas que nos permite generar, de forma mecánica, todas las secuencias $\vec{a}$ de números naturales $\vec{a} \in \mathbb{N}^n$ para las que sucede $R(\vec{a})$, de tal modo que cada una de estas secuencias aparezca exactamente una vez en la enumeración, y tras un número finito de pasos.

Y decimos que una relación es *"recursivamente enumerable"* (abreviadamente, *"r.e."*), cuando existe un procedimiento de enumeración para ella.

Un ejemplo interesante de relaciones recursivamente enumerables es el que estudiamos a continuación.

§ 4.6. Definición (Relaciones universales). Para cada número natural $n \geqslant 1$, la relación O_n será aquella relación n-aria que es cumplida por *todas* las secuencias n-arias de números naturales sin excepción. Es decir, aquella relación tal que:

$$O_n(\vec{a}) \qquad \text{para toda } \vec{a} \in \mathbb{N}^n$$

A la relación O_n para cada natural n, la llamamos *"relación universal n-aria"*.

§ 4.7. Práctica* (Orden lexicográfico).

(a) Clasificar las siguientes expresiones por orden alfabético, como si fuesen palabras en un diccionario, indicando el número de orden que corresponde a cada una:

"baaa"	"aaaa"	"abba"	"baba"
"abaa"	"abab"	"abbb"	"bbbb"
"bbba"	"aaab"	"baab"	"babb"
"aabb"	"aaba"	"bbaa"	"bbab"

(b) Clasificar las siguientes secuencias en un orden *lexicográfico* en el cual el "0" preceda al "1", al igual que la "a" precede a la "b" en nuestro alfabeto. Indicar el número de orden que corresponde a cada secuencia.

$1,0,0,0$	$0,0,0,0$	$0,1,1,0$	$1,0,1,0$
$0,1,0,0$	$0,1,0,1$	$0,1,1,1$	$1,1,1,1$
$1,1,1,0$	$0,0,0,1$	$1,0,0,1$	$1,0,1,1$
$0,0,1,1$	$0,0,1,0$	$1,1,0,0$	$1,1,0,1$

§ 4.8. Lema (Relaciones universales). *Todas las relaciones universales son recursivamente enumerables.*

Prueba. Para demostrar esto tomamos cualquier número natural $n \geqslant 1$, y a continuación especificamos cómo generar una enumeración de todas aquellas secuencias n-arias de naturales que cumplen la relación O_n, que son *todas*.

Empezamos por establecer un orden *lexicográfico* (o *"pseudo alfabético"*) entre los números naturales, conforme a su orden inmediato de menor a mayor: 0, 1, 2, 3, ... Así por ejemplo, el 0 y el 1 serían los correlatos de las letras "a" y "b" de nuestro alfabeto, respectivamente. A continuación vendría el 2, luego el 3, y así hasta el infinito.

Una vez hecho esto, consideramos al conjunto de secuencias n-arias cuya suma de componentes es igual a 0. Nada más que hay una, claro está, la secuencia

$$\overbrace{0,0,\ldots,0}^{n \text{ veces}}$$

A continuación, consideramos al conjunto de secuencias n-arias cuya suma de componentes es igual a 1. Y las generamos una a una, de acuerdo con el orden lexicográfico que acabamos de describir:

$$\overbrace{0,0,\ldots,0,1}^{n} \qquad \overbrace{0,0,\ldots,1,0}^{n} \qquad \ldots \qquad \overbrace{0,1,\ldots,0,0}^{n} \qquad \overbrace{1,0,\ldots,0,0}^{n}$$

A renglón seguido consideramos las secuencias n-arias cuya suma de componentes es 2. Y las vamos generando, siguiendo el mismo orden lexicográfico de colocación entre ellas.

Y así continuamos sucesivamente, con el conjunto de todas las secuencias n-arias que sumen 3, 4, etc. Estos conjuntos sucesivos se irán haciendo cada vez mayores, pero lo importante es que cada uno de ellos siempre será un conjunto finito, y que una vez completada la generación de cada uno, podemos pasar al siguiente y continuar la lista.

Cualquier secuencia n-aria dada de números naturales tendrá una determinada suma de componentes, por grande que sea, por lo que deberá aparecer en su lugar dentro del grupo que le corresponda. Y en definitiva, será alcanzada en nuestra enumeración, después de un número finito de pasos.

El resultado es, por tanto, un procedimiento mecánico para generar todas las secuencias n-arias de números naturales, sin excepción.

§ 4.9. Definición (Relación complementaria). Si R es una relación n-aria, entonces la *relación complementaria* de R es sencillamente aquella relación n-aria que se da para una secuencia $\vec{a}$ de números naturales justo en aquellos casos en que la relación R *no* se da para esa secuencia. Es decir: aquella relación que se da para una secuencia $\vec{a}$ de naturales justo en aquellos casos en que tenemos $\sim R(\vec{a})$.

Por ejemplo, la relación complementaria de la relación binaria *ser mayor que* es claramente la relación, también binaria, de *ser menor o igual que.*

§ 4.10. Práctica* (Relaciones complementarias). Mostrar que si R es una relación recursiva, entonces tanto R como la complementaria de R son relaciones recursivamente enumerables. Para ello, apoyarse en las relaciones universales.

§ 4.11. Observación (Problema 7). Algo más de ingenio hace falta para establecer el resultado recíproco: que si tanto R como su complementaria son r.e., entonces R tiene que ser recursiva. Ello es lo que se pide en el *Problema 7* del listado de *Problemas propuestos* (p. 12), y ahora es el momento apropiado para plantearse la respuesta a dicho problema.

§ 4.12. Definición (Función recursiva). Sea f una función n-aria. Un *procedimiento de computación* para f es un conjunto finito de instrucciones precisas que, aplicadas a cualquier secuencia $\vec{a} \in \mathbb{N}^n$, nos permiten calcular, de forma mecánica y en un número finito de pasos, el valor de $f(\vec{a})$.

Decimos que una función es *"recursiva"* cuando existe un procedimiento de computación para ella.

§ 4.13. Definición (Gráfico de una función). Sea f una función n-aria. El *gráfico* de f es aquella relación $(n+1)$-aria R tal que para cualesquiera $\vec{a} \in \mathbb{N}^n$ y $b \in \mathbb{N}$,

$$R(\vec{a}, b) \qquad \text{si y sólo si} \qquad f(\vec{a}) = b$$

Por consiguiente, el gráfico de una función n-aria es aquella relación que se da en una secuencia $(n+1)$-aria, cuando el último componente de la secuencia es precisamente el valor que toma dicha función al ser aplicada a los n componentes anteriores.

Esta idea puede parecer algo intrincada, pero si nos fijamos bien, no lo es tanto. El gráfico de la función monaria de sucesión, por ejemplo, que asigna a cada número natural n su sucesor $n+1$, no puede ser otra que la relación binaria *ser una unidad menor que*, que es justamente la relación que se da entre dos números naturales cuando el segundo es el sucesor del primero.

§ 4.14. Teorema (Funciones recursivas). *Para cualquier función f, las tres condiciones siguientes son equivalentes:*

(a) *f es recursiva.*

(b) *El gráfico de f es una relación recursiva.*

(c) *El gráfico de f es una relación r.e.*

Prueba. Sea f una función n-aria, y sea R su gráfico. Entonces obviamente, para cualesquiera $\vec{a} \in \mathbb{N}^n$ y $b \in \mathbb{N}$, tenemos

$$R(\vec{a}, b) \qquad \text{si y sólo si} \qquad f(\vec{a}) = b$$

Ahora vamos a demostrar, en primer lugar, que (a) implica (b). Supongamos que, en efecto, f es una función recursiva. Por tanto debe existir un procedimiento de computación para f, y entonces resulta muy fácil suplementar éste para obtener un procedimiento de decisión para R. En efecto, para cualesquiera $\vec{a} \in \mathbb{N}^n$ y $b \in \mathbb{N}$, basta con calcular el valor de $f(\vec{a})$, y según coincida o no con b, ya habremos determinado si $R(\vec{a}, b)$ sucede o no.

Que (b) implica (c) se sigue de inmediato por la *Práctica* § 4.10.

Y con eso sólo nos falta probar que a su vez (c) implica (a). Pues en efecto, suponiendo que R sea r.e., debe existir un procedimiento de enumeración para R. Puesto en marcha, dicho procedimiento generará, para cada $\vec{a} \in \mathbb{N}^n$, una secuencia $(n + 1)$-aria cuyos n primeros componentes serán los de $\vec{a}$. Es decir, una secuencia de la forma

$$\vec{a}, b$$

para algún $b \in \mathbb{N}$.

Nótese que debe aparecer exactamente *una* secuencia de esta forma, esto es, una secuencia cuyos n primeros componentes sean los de $\vec{a}$, ya que siendo f una función, el valor que tome para los argumentos de $\vec{a}$ ha de ser único (cf. § 1.41, p. 36).

Y siendo R su gráfico, sólo puede haber por tanto una secuencia de la forma $\vec{a}, b$ que cumpla R. Una vez que aparezca esta secuencia, sabremos obviamente que b era el valor de $f(\vec{a})$.

En definitiva, para calcular de forma mecánica el valor de $f(\vec{a})$, basta con poner en marcha el procedimiento de enumeración para R, esperar a que aparezca una secuencia cuyos primeros n componentes sean los de $\vec{a}$, que tendrá que aparecer más tarde o más temprano, después de un número finito de pasos, y entonces el valor de $f(\vec{a})$ será justamente el último componente de dicha secuencia.

Y de este modo, una vez demostrado que la condición (c) implica la (a), hemos "cerrado el círculo", por así decirlo. Esto es: hemos demostrado que las tres condiciones indicadas son estrictamente equivalentes.

El teorema de Matiyasevich

§ 4.15. Definición (Polinomios). Un *polinomio de coeficientes naturales* es una expresión algebraica consistente en una suma indicada, en la cual cada sumando consta de un producto de variables numéricas, elevadas a diversas potencias naturales, y multiplicadas por algún número, también natural, que constituye su *coeficiente*. A la vista de esta definición alguno puede asustarse, pero como vamos a ver enseguida por un ejemplo, en realidad se trata de expresiones que a todos nos deben resultar familiares.

Para abreviar, en este Manual diremos con frecuencia “polinomio” sin más, entendiendo que se trata siempre de polinomios de coeficientes naturales.

Además, si en un polinomio f aparecen un total de n variables numéricas, entonces diremos que f es un polinomio “n-ario”. En tal caso, y siendo $\vec{a} \in \mathbb{N}^n$ una secuencia n-aria de números naturales, pondremos

$$\text{“} f(\vec{a}) \text{”}$$

para denotar el valor que tome dicho polinomio cuando sustituimos las variables numéricas, en su orden de aparición, por los componentes de $\vec{a}$.

Un ejemplo sencillo de polinomio lo tenemos en:

$$h: \qquad 4x^3yz + 5y^2 + 7z + 1$$

Este polinomio consta, como vemos, de cuatro sumandos, en los cuales aparecen 3 variables numéricas distintas, x, y y z. Los coeficientes de los sumandos son 4, 5, 7 y 1, respectivamente. Y el valor que toma h para una secuencia concreta, digamos 1, 1, 2, por ejemplo, es fácil de calcular:

$$\begin{array}{rcccccccc} h(1,1,2) & = & (4 \cdot 1^3 \cdot 1 \cdot 2) & + & (5 \cdot 1^2) & + & (7 \cdot 2) & + & 1 \\ & = & 8 & + & 5 & + & 14 & + & 1 \\ & = & 28 & & & & & & \end{array}$$

§ 4.16. Práctica (Polinomios). Proponer un ejemplo de polinomio con 5 sumandos y 3 variables numéricas distintas, calculando su valor para la secuencia 1, 2, 3.

§ 4.17. Definición (Relaciones elementales). Una relación n-aria R es *elemental* cuando existen dos polinomios n-arios f y g, tales que para toda $\vec{a} \in \mathbb{N}^n$,

$$R(\vec{a}) \qquad \text{si y sólo si} \qquad f(\vec{a}) = g(\vec{a})$$

Un caso sumamente sencillo de relación elemental es la relación binaria *ser el cuadrado de*, que se puede representar, por ejemplo, mediante los siguientes polinomios, ambos binarios:

$$h_1: \qquad xy + x \qquad\qquad\qquad h_2: \qquad xy + y^2$$

En efecto, para cualquier pareja de naturales a, b, está claro que $ab + a$ será igual a $ab + b^2$ si y sólo si $a = b^2$, con lo que

$$h_1(a,b) = h_2(a,b) \qquad \text{si y sólo si} \qquad a \text{ es el cuadrado de } b$$

§ 4.18. Lema (Relaciones elementales). *Cualquier relación elemental es recursiva.*

Prueba. Sea R cualquier relación elemental n-aria. Entonces por definición han de existir dos polinomios n-arios f y g tales que para cualquier $\vec{a} \in \mathbb{N}^n$,

$$R(\vec{a}) \qquad \text{si y sólo si} \qquad f(\vec{a}) = g(\vec{a})$$

Ahora bien, dada cualquier $\vec{a} \in \mathbb{N}^n$, las reglas de la aritmética elemental nos permiten calcular en un número finito de pasos los valores de $f(\vec{a})$ y $g(\vec{a})$, y por tanto decidir si son o no iguales. Y con ello habremos determinado si $R(\vec{a})$ sucede o no.

§ 4.19. Definición (Relaciones diofánticas). Por su parte, una relación n-aria Q es *diofántica* cuando existe alguna relación elemental R, tal que para toda $\vec{a} \in \mathbb{N}^n$, tenemos

$$Q(\vec{a}) \qquad \text{si y sólo si} \qquad \text{existe una secuencia } \vec{b} \in \mathbb{N}^m \text{ tal que } R(\vec{a}, \vec{b})$$

En otras palabras: que para que Q sea diofántica tiene que suceder que Q se dé para una secuencia de naturales $\vec{a} \in \mathbb{N}^n$ si y sólo si existe una segunda secuencia $\vec{b} \in \mathbb{N}^m$ de tal modo que la relación elemental R se dé a su vez para la secuencia continua $\vec{a}, \vec{b}$. Es decir, de tal modo que:

$$R(a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_m)$$

Naturalmente, la relación elemental R que está vinculada a Q será en este caso una relación $(n+m)$-aria. Donde por cierto, m y n son números naturales que no tienen por qué ser iguales, naturalmente.

En fin, un ejemplo muy sencillito de relación diofántica, que nos puede ayudar a familiarizarnos con esta definición, lo tenemos en la propiedad (relación monaria) de *ser un cuadrado perfecto.* Como sabemos, se llaman "cuadrados perfectos" a aquellos números naturales que resultan de elevar otro al cuadrado, como el 4, el 9, el 25 ó el 49.

Pues bien, podemos decir que $a \in \mathbb{N}$ es un cuadrado perfecto si y sólo si existe un $b \in \mathbb{N}$ tal que a es el cuadrado de b. Y dado que la relación (binaria) *ser el cuadrado de* es una relación elemental, como ya hemos visto, queda claro por lo tanto que la relación (monaria) *ser un cuadrado perfecto* es a su vez una relación diofántica.

§ 4.20. Teorema (Relaciones diofánticas). *Cualquier relación diofántica es r.e.*

Prueba. Sea Q cualquier relación diofántica n-aria. Por definición ha de existir una relación elemental R tal que para cualquier $\vec{a} \in \mathbb{N}^n$,

$$Q(\vec{a}) \qquad \text{si y sólo si} \qquad \text{existe una secuencia } \vec{b} \in \mathbb{N}^m \text{ tal que } R(\vec{a}, \vec{b})$$

Entonces, puesto que R es elemental, por el *Lema* § 4.18 es recursiva, y por tanto ha de ser también r.e. Consideremos entonces un procedimiento de enumeración para R, y modifiquemos sus instrucciones de forma que cada vez que se genere una secuencia $\vec{a}, b_1, b_2, \ldots, b_m$, los últimos m componentes sean borrados, dejando sólo a $\vec{a}$.

Y de esta sencilla forma, si se van omitiendo las repeticiones que pudieran ocurrir, obtendremos una enumeración exacta y completa de todas las secuencias $\vec{a} \in \mathbb{N}^n$ tales que $Q(\vec{a})$.

§ 4.21. Teorema (Teorema de Matiyasevich). *Cualquier relación r.e. es diofántica.*

Prueba. Este importante teorema, converso del anterior, tiene una larga y compleja prueba, cuyo último tramo fue completado en 1970 por el matemático ruso Yuri Matiyasevich. Aunque cronológicamente es bastante posterior a la demostración de los teoremas limitativos que vamos a ver después, y por lo tanto *no* es estrictamente necesario para probar aquellos, nosotros nos apoyaremos el él, omitiendo su prueba, a fin de facilitar nuestro camino para la obtención de dichos resultados.

MÓDULO 5

El teorema de Tarski

Aritmeticidad

§ 5.1. Convención (Sustituciones por numerales). Sea $\vec{a} \in \mathbb{N}^n$ cualquier secuencia n-aria de números naturales. Entonces, naturalmente, sus componentes

$$a_1 \quad a_2 \quad \ldots \quad a_n$$

serán todos ellos números naturales.

Consideremos ahora los numerales que corresponden a dichos componentes, es decir, los numerales que tienen a esos números naturales como subíndices:

$$\underline{s_{a_1}} \quad \underline{s_{a_2}} \quad \ldots \quad \underline{s_{a_n}}$$

Por ejemplo, si a_3 resultara ser el número 27, entonces $\underline{s_{a_3}}$ sería el numeral $\underline{s_{27}}$, es decir, aquel numeral que consiste en 27 apariciones del símbolo de sucesión $\underline{s}$, seguidas por la constante $\underline{0}$.

Pues bien. A nosotros nos va a interesar a partir de ahora el resultado de efectuar, en una fórmula cualquiera $\underline{\alpha}$, las sustituciones sucesivas

$$[\,\underline{v_1 \,/\, s_{a_1}}\,] \quad [\,\underline{v_2 \,/\, s_{a_2}}\,] \quad \ldots \quad [\,\underline{v_n \,/\, s_{a_n}}\,]$$

Es decir: la sustitución de la variable $\underline{v_1}$ de $\mathscr{A}$ por el numeral $\underline{s_{a_1}}$, a continuación, una vez efectuada ésta, la sustitución de la variable $\underline{v_2}$ por el numeral $\underline{s_{a_2}}$, y así sucesivamente hasta sustituir $\underline{v_n}$ por $\underline{s_{a_n}}$.

Al resultado de efectuar completamente esta operación lo vamos a denotar de forma abreviada poniendo

$$\text{“}\,\underline{\alpha[s_{\vec{a}}]}\,\text{”}$$

(que puede leerse sencillamente “$\alpha - s - a$ *vector*”).

§ 5.2. Observación (Sustituciones por numerales). Está claro que si $\underline{\alpha}$ tiene variables libres posteriores a $\underline{v_n}$, entonces la fórmula sustituida $\underline{\alpha[s_{\vec{a}}]}$ las seguirá teniendo, ya que la sustitución indicada sólo afecta a las n primeras variables de $\mathscr{A}$.

Sin embargo, si por el contrario, $\underline{\alpha}$ no tiene variables libres posteriores a $\underline{v_n}$, es decir, si $\underline{\alpha} \in \Upsilon_n$, entonces, como todas las variables hasta $\underline{v_n}$ son sustituidas por numerales, que son términos cerrados, la fórmula resultante $\underline{\alpha[s_{\vec{a}}]}$ será necesariamente una oración. En otras palabras:

$$\text{si} \qquad \vec{a} \in \mathbb{N}^n \qquad \text{y} \qquad \underline{\alpha} \in \Upsilon_n \qquad \text{entonces} \qquad \underline{\alpha[s_{\vec{a}}]} \text{ es una oración}$$

Por último, hemos de tener presente el caso en que la secuencia en cuestión consiste en un único número natural $a \in \mathbb{N}$, esto es, lo que llamamos en su momento una "secuencia monaria" (cf. § 1.38, p. 35). Entonces, aplicando el razonamiento anterior se sigue de inmediato que

$$\text{si} \qquad a \in \mathbb{N} \qquad \text{y} \qquad \underline{\alpha} \in \Upsilon_1 \qquad \text{entonces} \qquad \underline{\alpha[s_a]} \text{ es una oración}$$

§ 5.3. Definición (Representabilidad). Sean

- R una relación n-aria
- Δ una teoría de $\mathscr{A}$
- $\underline{\alpha}$ una fórmula de $\mathscr{A}$ tal que $\underline{\alpha} \in \Upsilon_n$

Entonces:

1. $\underline{\alpha}$ *representa débilmente* a R en la teoría Δ cuando para cualquier $\vec{a} \in \mathbb{N}^n$,

$$\begin{cases} \text{si} \quad R(\vec{a}) \quad \text{entonces} \quad \underline{\alpha[s_{\vec{a}}]} \in \Delta \\ \text{si} \quad \sim R(\vec{a}) \quad \text{entonces} \quad \underline{\alpha[s_{\vec{a}}]} \notin \Delta \end{cases}$$

2. $\underline{\alpha}$ *representa fuertemente* a R en la teoría Δ cuando para cualquier $\vec{a} \in \mathbb{N}^n$,

$$\begin{cases} \text{si} \quad R(\vec{a}) \quad \text{entonces} \quad \underline{\alpha[s_{\vec{a}}]} \in \Delta \\ \text{si} \quad \sim R(\vec{a}) \quad \text{entonces} \quad \underline{\neg\alpha[s_{\vec{a}}]} \in \Delta \end{cases}$$

Decimos que una relación es *"débilmente representable"* en una teoría cuando existe alguna fórmula que la representa débilmente en esa teoría. Y decimos que es *"fuertemente representable"* cuando existe alguna fórmula que la representa fuertemente en esa teoría.

§ 5.4. Observación (Representabilidad). Como ya hemos visto, si $\underline{\alpha} \in \Upsilon_n$, la fórmula sustituida $\underline{\alpha[s_{\vec{a}}]}$ será una oración, y es por ello que tiene sentido plantearse si pertenece o no a una determinada teoría, Δ. Y exactamente lo mismo cabe decir de su negación, $\underline{\neg\alpha[s_{\vec{a}}]}$.

Por lo demás, nótese que si de hecho sucede $R(\vec{a})$, entonces la oración $\underline{\alpha[s_{\vec{a}}]}$ habrá de pertenecer a la correspondiente teoría Δ en ambos tipos de representación. Y si sucede que $\sim R(\vec{a})$, es decir, si *no* sucede que $R(\vec{a})$, entonces, en el caso de la representación débil, la oración $\underline{\alpha[s_{\vec{a}}]}$ no podrá pertenecer a Δ. Y en el caso de la representación fuerte, será la oración $\underline{\neg\alpha[s_{\vec{a}}]}$ la que sí deba pertenecer a Δ.

Si la teoría en cuestión, es *consistente*, entonces la representación fuerte implica la débil. En efecto, ello es así porque si $\underline{\neg\alpha[s_{\vec{a}}]} \in \Delta$ y Δ es consistente, entonces naturalmente $\underline{\alpha[s_{\vec{a}}]} \notin \Delta$. Es en este sentido que la representación fuerte es "más fuerte" que la débil. Si la teoría es inconsistente, por el contrario, dicha implicación no se da. Pero es poco relevante cómo sea la representación en una teoría inconsistente, en cualquier caso.

Por último, adviértase también que para cualquier fórmula $\underline{\alpha} \in \Upsilon_n$ y teoría Δ, siempre habrá una relación que $\underline{\alpha}$ represente débilmente en Δ, a saber: aquella relación que es cumplida por una secuencia de naturales $\vec{a} \in \mathbb{N}^n$ si y sólo si $\underline{\alpha[s_{\vec{a}}]} \in \Delta$. Sin embargo, es posible que $\underline{\alpha}$ no represente fuertemente ninguna relación en Δ, que será lo que suceda en aquellos casos en los que exista algún $\vec{a} \in \mathbb{N}^n$ tal que ni $\underline{\alpha[s_{\vec{a}}]}$ ni $\underline{\neg\alpha[s_{\vec{a}}]}$ pertenezcan a Δ.

§ 5.5. Observación (Problema 8). Este es el momento apropiado para plantearse la respuesta al *Problema 8* del listado de *Problemas propuestos* (p. 13).

§ 5.6. Observación (Representabilidad en Ω). Es evidente que para cualquier oración $\underline{\beta}$ de $\mathscr{A}$, o bien $\underline{\beta} \in \Omega$, o bien $\underline{\neg\beta} \in \Omega$, ya que necesariamente una de las dos ha de ser verdadera en el modelo $\mathscr{N}$.

Por lo tanto, para cualquier $\underline{\alpha} \in \Upsilon_n$, como la correspondiente fórmula sustituida $\underline{\alpha[s_{\vec{a}}]}$ es siempre una oración, según lo que ya hemos visto, tendremos que o bien $\underline{\alpha[s_{\vec{a}}]} \in \Omega$ o bien $\underline{\neg\alpha[s_{\vec{a}}]} \in \Omega$.

Por consiguiente, cada fórmula $\underline{\alpha} \in \Upsilon_n$ representa una única relación n-aria en Ω, tanto débil como fuertemente. Y ello justifica la siguiente definición.

§ 5.7. Definición (Aritmeticidad). Una relación n-aria R es *aritmética* cuando es representable en Ω, es decir: cuando existe una fórmula $\underline{\alpha} \in \Upsilon_n$ que la representa en Ω.

De acuerdo con la observación precedente, en la teoría Ω la representación débil es equivalente a la fuerte, por lo cual, a la hora de establecer la aritmeticidad de una relación, podemos utilizar indistintamente cualquiera de los dos tipos de representación.

§ 5.8. Definición (Función aritmética). Una función es *aritmética* cuando su gráfico es una relación aritmética.

§ 5.9. Práctica* (Relaciones aritméticas).

(a) Determinar las relaciones representadas en Ω por las siguientes cuatro fórmulas de $\mathscr{A}$:

$$\underline{v_1 = v_2 \times s_2} \qquad\qquad \underline{\exists v_3\,(\,v_1 = v_2 + v_3\,)}$$

$$\underline{\exists v_2\,(\,v_1 = v_2 \times s_2\,)} \qquad\qquad \underline{\exists v_3\,(\,v_3 \neq 0 \;\wedge\; v_1 = v_2 + v_3\,)}$$

(b) Indicar una fórmula de $\mathscr{A}$ que represente en Ω la relación binaria *ser el cuadrado de.*

(c) Indicar una fórmula de $\mathscr{A}$ que represente en Ω la propiedad de *ser un cuadrado perfecto.*

(d) Indicar una fórmula de $\mathscr{A}$ que represente en Ω el gráfico de la función de sucesión.

§ 5.10. Lema (Relaciones aritméticas). *Cualquier relación elemental es aritmética.*

Prueba. Siendo R una relación elemental n-aria, por definición tendrán que existir dos polinomios n-arios f y g tales que para toda $\vec{a} \in \mathbb{N}^n$,

$$R(\vec{a}) \qquad \text{si y sólo si} \qquad f(\vec{a}) = g(\vec{a})$$

Pues bien: para componer una fórmula de $\mathscr{A}$ que represente a R en Ω basta con trasladar esta igualdad entre polinomios al lenguaje formal $\mathscr{A}$, de la manera más obvia:

- Reemplazando las variables numéricas de los polinomios, en su orden de aparición en la igualdad señalada, por las variables de $\mathscr{A}$, $\underline{v_1}, \underline{v_2}, \ldots, \underline{v_n}$, en ese orden.
- Reemplazando los coeficientes numéricos de cada polinomio, por sus correspondientes numerales de $\mathscr{A}$.
- Reemplazando las sumas y productos indicados, por los símbolos de adición y multiplicación de $\mathscr{A}$, respectivamente.
- Representando las potencias en los polinomios mediante multiplicaciones repetidas.
- Y finalmente, reemplazando la igualdad entre los dos polinomios por el símbolo de igualdad de $\mathscr{A}$.

Es evidente que la fórmula resultante, digamos "$\underline{\beta}$", será verdadera bajo una evaluación basada en $\mathscr{N}$ si y sólo si los valores que asigna a las variables $\underline{v_1}, \underline{v_2}, \ldots, \underline{v_n}$, conforman una secuencia $\vec{a} \in \mathbb{N}^n$ que cumple $f(\vec{a}) = g(\vec{a})$, y por consiguiente, que cumple $R(\vec{a})$.

Por lo tanto, como la correspondiente oración $\underline{\beta[s_{\vec{a}}]}$ resulta precisamente de substituir las variables $\underline{v_1}, \underline{v_2}, \ldots, \underline{v_n}$ por los numerales $\underline{s_{a_1}}, \underline{s_{a_2}}, \ldots, \underline{s_{a_n}}$, resulta patente que $\underline{\beta[s_{\vec{a}}]}$ será verdadera en $\mathscr{N}$ si y sólo si sucede que $R(\vec{a})$. Es decir: $\underline{\beta[s_{\vec{a}}]} \in \Omega$ si y sólo si $R(\vec{a})$.

Y de este modo concluimos que, en efecto, $\underline{\beta}$ representa R en Ω, esto es: R es aritmética.

§ 5.11. Práctica* (Más relaciones aritméticas). Indicar una fórmula de $\mathscr{A}$ que represente en Ω la relación ternaria S definida por

$$S(x, y, z) \qquad \text{si y sólo si} \qquad 4x^3yz + 5y^2 + 7z + 1 = 3xyz + 2z$$

§ 5.12. Teorema (Relaciones aritméticas). *Cualquier relación diofántica es aritmética.*

Prueba. Sea Q una relación diofántica n-aria. Entonces, por definición ha de existir una relación elemental R tal que para cualquier $\vec{a} \in \mathbb{N}^n$,

$$Q(\vec{a}) \qquad \text{si y sólo si} \qquad \text{existe una secuencia } \vec{b} \in \mathbb{N}^m \text{ tal que } R(\vec{a}, \vec{b})$$

Puesto que la relación R es elemental, por el lema precedente existe una fórmula $\underline{\beta}$ que la representa en Ω. Como R es $(n+m)$-aria, tenemos que en este caso, $\underline{\beta} \in \Upsilon_{n+m}$. Pues bien, para componer la fórmula $\underline{\alpha}$ de $\mathscr{A}$ que representa a Q en Ω basta con poner sencillamente

$$\underline{\exists\, v_{n+1}\, v_{n+2}\, \ldots\, v_{n+m}\;\; \beta}$$

Es claro que la fórmula así construida $\underline{\alpha} \in \Upsilon_n$, ya que las variables $\underline{v_{n+1}}, \underline{v_{n+2}}, \ldots, \underline{v_{n+m}}$, han quedado ligadas por la cuantificación inicial. Además, es inmediato ver que para cualquier $\vec{a} \in \mathbb{N}^n$, la oración $\underline{\alpha[s_{\vec{a}}]}$, esto es, la oración

$$\underline{\exists\, v_{n+1}\, v_{n+2}\, \ldots\, v_{n+m}\;\; \beta[s_{\vec{a}}]}$$

será verdadera en $\mathscr{N}$ si y sólo si existe una secuencia $\vec{b} \in \mathbb{N}^m$ tal que la oración correspondiente $\underline{\beta[s_{\vec{a},\vec{b}}]}$ sea verdadera en $\mathscr{N}$. Y como $\underline{\beta}$ representa la relación R, esto significa a su vez que $R(\vec{a}, \vec{b})$.

En otras palabras: que $\underline{\alpha[s_{\vec{a}}]}$ será verdadera en $\mathscr{N}$ si y sólo si existe una secuencia $\vec{b} \in \mathbb{N}^m$ tal que $R(\vec{a}, \vec{b})$.

Y de este modo concluimos que $\underline{\alpha}$ representa a Q en Ω, y por consiguiente, que Q es aritmética.

§ 5.13. Corolario (Relaciones aritméticas). *Cualquier relación r.e. es aritmética.*

Prueba. Inmediata por el teorema precedente y el teorema de Matiyasevich (§ 4.21, p. 95).

§ 5.14. Corolario (Funciones aritméticas). *Cualquier función recursiva es aritmética.*

Prueba. Inmediata por el corolario precedente y § 4.14 (p. 92).

El teorema de Tarski

§ 5.15. Definición (Codificación). A cada símbolo primitivo de $\mathscr{A}$ le asignamos un *código* o *número de Gödel* (representado por el símbolo "$\sharp$", que se lee *"código de"*), mediante la siguiente tabla de correspondencias:

$\sharp\,\underline{0}$	10	$\sharp\,\underline{=}$	10^5	$\sharp\,\underline{v_1}$	10^9
$\sharp\,\underline{s}$	100	$\sharp\,\underline{\neg}$	10^6	$\sharp\,\underline{v_2}$	10^{10}
$\sharp\,\underline{+}$	1000	$\sharp\,\underline{\rightarrow}$	10^7	$\sharp\,\underline{v_3}$	10^{11}
$\sharp\,\underline{\times}$	10000	$\sharp\,\underline{\forall}$	10^8		$\ldots$

Asimismo, a cada cadena de símbolos de $\mathscr{A}$ le asignamos también un *código* o *número de Gödel*, que consiste sencillamente en la concatenación de los códigos de cada uno de los símbolos que aparecen en la cadena.

Para calcular el código de una cadena de $\mathscr{A}$, habrá que tener en cuenta, naturalmente, el orden original en el que se colocan los símbolos en este lenguaje formal, y no el de nuestras abreviaturas y convenciones. Por ejemplo:

$$\begin{aligned} \sharp\,\underline{0 = v_1} &= \sharp\,\underline{= 0 v_1} \\ &= \underbrace{100000}_{5\text{ ceros}}\,10\,\underbrace{1000000000}_{9\text{ ceros}} \end{aligned}$$

$$\begin{aligned} \sharp\,\underline{0 \neq s_2} &= \sharp\,\underline{\neg = 0 s_2} \\ &= \sharp\,\underline{\neg = 0 s s 0} \\ &= \underbrace{1000000}_{6\text{ ceros}}\,\underbrace{100000}_{5\text{ ceros}}\,1010010010 \end{aligned}$$

§ 5.16. Práctica⋆ (Codificación).

(a) Calcular $\sharp\,\underline{0 + 0 = 0}$.

(b) Indicar a qué fórmula de $\mathscr{A}$ corresponde el código $\underbrace{100000}_{5\text{ ceros}}\,\underbrace{1000}_{3\text{ ceros}}\,101010$.

§ 5.17. Teorema (Códigos de fórmulas). *Las propiedades de* ser el código de una cadena*,* ser el código de un término*,* ser el código de una fórmula *y* ser el código de una oración*, son recursivas.*

Prueba. Para determinar si un número natural cualquiera es el código de una cadena, basta con examinarlo dígito a dígito, en su representación ordinaria en notación decimal, y comprobar:

- que consta únicamente de los dígitos “0” y “1”
- que empieza por “1” y termina por “0”

- que el "1" nunca aparece dos veces seguidas.

Y se sigue de inmediato de la *Definición* § 5.15, que un número natural es el código de una cadena si y sólo si cumple con estas tres condiciones.

Una vez hecho esto, resulta sencillo identificar los símbolos de $\mathscr{A}$ que componen la cadena en cuestión, utilizando nuestra tabla de codificación, y teniendo en cuenta que el código de cada símbolo comienza siempre por "1".

Y por último, resulta a su vez inmediato verificar si se trata del código de un término, fórmula u oración, simplemente aplicando las definiciones correspondientes a estas tres nociones.

§ 5.18. Corolario (Códigos de fórmulas). *Las propiedades de* ser el código de una cadena*,* ser el código de un término*,* ser el código de una fórmula *y* ser el código de una oración*, son aritméticas.*

Prueba. Inmediata por el teorema precedente, la *Práctica* § 4.10 (p. 92) y el *Corolario* § 5.13.

§ 5.19. Observación (Nociones sintácticas). De acuerdo con el corolario precedente, en $\mathscr{A}$ tiene que existir, por ejemplo, una fórmula $\underline{\alpha}$ que represente en Ω la propiedad de ser el código de un término. Es decir: una fórmula $\underline{\alpha}$ tal que para todo $a \in \mathbb{N}$,

$$a \text{ es el código de un término} \qquad \text{si y sólo si} \qquad \underline{\alpha[s_a]} \in \Omega$$

He aquí una capacidad sorprendente del método de codificación: consigue expresar mediante una fórmula de $\mathscr{A}$, y por referencia a la teoría Ω, cierta noción básica relativa al propio lenguaje $\mathscr{A}$, como es la de *ser un término*. Bien es verdad que se trata de un procedimiento indirecto y artificioso, en el cual la propiedad representada es en realidad la de *ser el código de un número natural que corresponde a un término de* $\mathscr{A}$. Pero del código se puede extraer inmediatamente el término a que corresponde, a partir de nuestra tabla de codificación. Y lo asombroso en cualquier caso es que exista una fórmula que presente dicha correlación.

Lo mismo cabe decir, naturalmente, del resto de nociones mencionadas en el *Teorema* § 5.17: las de *cadena*, *fórmula* y *oración*, que también son representables en Ω.

A estas nociones y otras muchas las llamamos técnicamente *"nociones sintácticas"*, porque se pueden definir sin hacer referencia a ninguna interpretación del lenguaje formal en cuestión. Se trata, por tanto, de nociones *puramente formales*. Pues bien, como es fácil imaginar, todas las nociones de este tipo son recursivas, o al menos r.e., y por tanto son todas ellas igualmente representables en Ω, es decir: son también propiedades o relaciones aritméticas.

§ 5.20. Definición (La función diagonal). Sea a un número natural cualquiera. Para definir el valor que toma la *función diagonal* en a, tenemos que considerar dos casos. Si resulta que a es el código de una fórmula $\underline{\alpha}$ de nuestro lenguaje formal, entonces la *función diagonal* (*diag*) es aquella función que asigna a a el código de la correspondiente fórmula sustituida $\underline{\alpha[s_a]}$, esto es: $diag\,(a) = \sharp\, \underline{\alpha[s_a]}$

Y si por el contrario, resulta que a no coincide con el código de ninguna fórmula de nuestro lenguaje, entonces la *función diagonal* lo deja como está, es decir: $diag\,(a) = a$.

En resumidas cuentas:

$$diag\,(a) = \begin{cases} \sharp\,\underline{\alpha[s_a]} & \text{si } a \text{ es el código de una fórmula } \underline{\alpha} \\ \\ a & \text{si } a \text{ no es código de ninguna fórmula} \end{cases}$$

Según esto, por ejemplo, $diag\,(23) = 23$, ya que el número 23 claramente no es el código de ninguna fórmula. Sin embargo, si tomamos por ejemplo el número

$$b = 100000101000000000$$

que es el código de la fórmula $\underline{0 = v_1}$, entonces

$$diag\,(b)$$

será a su vez el código de la fórmula resultante de efectuar la sustitución $\underline{0 = v_1\,[v_1/s_b]}$. Es decir,

$$\underline{0 = v_1\,[v_1/s_b]} \qquad = \qquad \underline{0 = s_b} \qquad = \qquad \underline{0 = s_{100000101000000000}}$$

cuyo código podemos calcular con facilidad, obteniendo un numerito ciertamente algo grande:

$$\begin{aligned} \sharp\,\underline{0 = s_{100000101000000000}} \quad &= \quad \sharp\,\underline{= 0 s_{100000101000000000}} \\ &= \quad \sharp\,\underline{= 0\,\underbrace{ss\;\ldots\;s}_{b \text{ veces}}\,0} \qquad = \quad 10000010\underbrace{\widehat{100}\;\;\widehat{100}\;\;\ldots\;\;\widehat{100}}_{b \text{ veces}}10 \end{aligned}$$

§ 5.21. Observación (La función diagonal). En fin, si hay algo que podemos sacar en claro de esta caprichosa función, es que la diagonal del código de una fórmula $\underline{\alpha}$ será siempre el código de $\underline{\alpha[s_{\sharp\alpha}]}$. Esto es:

$$\text{para cualquier fórmula } \underline{\alpha} \text{ de } \mathscr{A}, \qquad diag\,(\sharp\,\underline{\alpha}) \quad = \quad \sharp\,\underline{\alpha[s_{\sharp\alpha}]}$$

§ 5.22. Práctica* (La función diagonal). Probar que la función diagonal es recursiva, y por lo tanto aritmética.

§ 5.23. Práctica* (Propiedades compuestas). Sean f una función monaria, y P y Q dos propiedades tales que para todo $a \in \mathbb{N}$,

$$P(a) \qquad \text{si y sólo si} \qquad Q(f(a))$$

Es decir: que un número natural a cumple la propiedad P exactamente en aquellos casos en que el número resultante de efectuar la operación $f(a)$ cumple la propiedad Q.

En estas condiciones, se cumple que:

(a) Si Q y f son aritméticas, también han de serlo tanto P como la propiedad complementaria de P. (Para probar esto hay que componer una nueva fórmula a partir de las representantes de Q y del gráfico de f.)

(b) Si Q y f son recursivas, también han de serlo tanto P como la propiedad complementaria de P.

(c) Si f es recursiva y Q es r.e., entonces P ha de ser también forzosamente r.e. (Para probar esto hay que componer un procedimiento de enumeración para P, alternando cíclicamente los de Q y del gráfico de f.)

Proporcionar la prueba para una de estas tres condiciones (a elección).

§ 5.24. Definición (La propiedad ***Or*****).** Sea Φ un conjunto de oraciones. Entonces Or_Φ es la propiedad numérica de *ser el código de una oración del conjunto* Φ. Es decir: aquella propiedad tal que para cualquier $a \in \mathbb{N}$,

$$Or_\Phi\,(a) \qquad \text{si y sólo si} \qquad a \text{ es el código de una oración perteneciente a } \Phi$$

§ 5.25. Teorema (Teorema de Tarski). *La propiedad de* ser el código de una oración verdadera en $\mathscr{N}$ *no es aritmética.*

Prueba. Puesto que el conjunto de todas las oraciones verdaderas en $\mathscr{N}$ es Ω, evidentemente la propiedad de *ser el código de una oración verdadera en* $\mathscr{N}$ es idéntica a la propiedad Or_Ω. Por lo tanto lo que dice este teorema es, puesto de otra forma, que la propiedad Or_Ω no es aritmética.

Supongamos ahora que Or_Ω fuese aritmética, y sea P aquella propiedad numérica tal que para todo $a \in \mathbb{N}$,

$$P(a) \qquad \text{si y sólo si} \qquad Or_\Omega\,(diag\,(a))$$

Es decir: que un número natural a cumple la propiedad P si y sólo si el número resultante de efectuar la operación $diag\,(a)$ cumple la propiedad Or_Ω.

De acuerdo con la *Práctica* § 5.23(a), suponiendo que Or_Ω fuese aritmética, también habrían de serlo forzosamente tanto la propiedad P como la propiedad complementaria de P. Por lo tanto, en particular, habrá de existir una fórmula $\underline{\alpha} \in \Upsilon_1$ que represente en Ω a la propiedad complementaria de P. Es decir, una fórmula $\underline{\alpha}$ tal que para cualquier $a \in \mathbb{N}$:

$$\underline{\alpha[s_a]} \in \Omega \qquad \text{si y sólo si} \qquad \sim P(a)$$

Ahora bien. Dicha fórmula $\underline{\alpha}$ es al fin y al cabo una cadena de símbolos de nuestro lenguaje formal, y tendrá por lo tanto un código, $\sharp\,\underline{\alpha}$, que será un determinado número natural.

Así pues, podemos considerar la correspondiente sustitución $\underline{\alpha[s_{\sharp\alpha}]}$. Y ello nos lleva a efectuar el razonamiento siguiente:

$$\underline{\alpha[s_{\sharp\alpha}]} \in \Omega \qquad \text{si y sólo si} \qquad \sim P(\sharp\,\underline{\alpha})$$

y a su vez,

$$\sim P(\sharp\,\underline{\alpha}) \qquad \text{si y sólo si} \qquad \sim Or_\Omega\,(diag\,(\sharp\,\underline{\alpha})) \qquad \text{(por definición de } P)$$

" $\sim Or_\Omega\,(\sharp\,\underline{\alpha[s_{\sharp\alpha}]})$ (por la *Observación* § 5.21)

" $\sharp\,\underline{\alpha[s_{\sharp\alpha}]}$ no es el código de una oración de Ω

" $\underline{\alpha[s_{\sharp\alpha}]} \notin \Omega$

En definitiva, hemos llegado a la conclusión de que $\underline{\alpha[s_{\sharp\alpha}]} \in \Omega$ si y sólo si $\underline{\alpha[s_{\sharp\alpha}]} \notin \Omega$, lo cual es imposible. En consecuencia, lo que hemos demostrado es que la hipotética fórmula $\underline{\alpha}$ no puede existir. Y por lo tanto, ni la propiedad complementaria de P puede ser aritmética, ni pueden serlo tampoco, por consiguiente, las propiedades P u Or_Ω.

§ 5.26. Observación (El teorema de Tarski). El teorema precedente se atribuye generalmente al gran lógico polaco Alfred Tarski, con fecha de 1933, al margen de ciertas disputas históricas que ha suscitado, y que a nosotros no nos interesan en absoluto.

A la vista de la prueba del teorema, no es de extrañar que la fórmula $\underline{\alpha}$ allí descrita no pueda existir, ya que para cada $a \in \mathbb{N}$, la oración $\underline{\alpha[s_a]}$ pertenecería a Ω, y por tanto, sería verdadera en $\mathscr{N}$, si y sólo si *diag* (a) *no* fuese el código de una oración de Ω, es decir, no fuese el código de una oración verdadera en $\mathscr{N}$. Así pues, la oración $\underline{\alpha[s_a]}$ vendría a expresar indirectamente, a través del método de codificación y por referencia a la teoría Ω, el hecho de que *diag* (a) no es el código de una oración verdadera en $\mathscr{N}$.

Por lo tanto, tomando el propio código de $\underline{\alpha}$, tenemos que la oración resultante $\underline{\alpha[s_{\sharp\alpha}]}$ expresaría indirectamente, a través del método de codificación, que *diag* $(\sharp\,\underline{\alpha})$ no es el código de una oración verdadera en $\mathscr{N}$, es decir: que la propia $\underline{\alpha[s_{\sharp\,\alpha}]}$ es falsa en $\mathscr{N}$.

La oración $\underline{\alpha[s_{\sharp\alpha}]}$ vendría a ser algo así como una reproducción de la *paradoja del mentiroso* en nuestro lenguaje formal $\mathscr{A}$: una oración que vendría a decir indirectamente, a través del método de codificación, algo así como "yo soy falsa". O al menos, "yo soy falsa en la interpretación estándar $\mathscr{N}$". Y lo que se prueba en el teorema de Tarski es que dicha fórmula no existe, de modo que nuestro lenguaje formal no se encuentra afectado por esta paradoja.

§ 5.27. Observación (Nociones semánticas). Por oposición a las nociones sintácticas, llamamos *"nociones semánticas"* a todas aquellas nociones relativas a las *interpretaciones* de un lenguaje formal. Pues bien, según vemos ahora, al contrario de lo que ocurría con las nociones sintácticas, una noción semántica fundamental relativa a nuestro lenguaje formal $\mathscr{A}$, como es la de *ser verdadero en la interpretación estándar* $\mathscr{N}$, resulta que *no* se puede representar mediante una fórmula de $\mathscr{A}$. Es decir: no es expresable mediante el lenguaje $\mathscr{A}$ y por referencia a la teoría Ω, ni siquiera mediante la estrategia indirecta que supone el método de codificación.

Ello se suele indicar condensadamente diciendo que *"la verdad no es una propiedad aritmética"*. E implica, claro está, que el lenguaje $\mathscr{A}$ es incapaz de funcionar como su propio metalenguaje, ya que resulta imposible representar mediante fórmulas suyas las nociones semánticas más básicas.

Y más en general, no es difícil ver que un argumento parecido al usado en la prueba del teorema de Tarski, se aplicará también a cualquier otro lenguaje formal que sea capaz de expresar, por referencia a alguna teoría, las nociones fundamentales relativas a su propia sintaxis, y en particular, la función análoga a la función diagonal, para ese lenguaje.

En lo que queda de Módulo exploraremos una segunda consecuencia limitativa del teorema de Tarski que es de capital importancia.

Teorías recursivamente axiomatizables

§ 5.28. Definición (Conjuntos de oraciones recursivos y recursivamente enumerables). Sea Φ un conjunto de oraciones de $\mathscr{A}$.

1. Φ es *recursivo* cuando la correspondiente propiedad Or_Φ, de *ser el código de una oración perteneciente a* Φ, es una propiedad recursiva.

2. Φ es *recursivamente enumerable* (abreviadamente, *"r.e."*), cuando la correspondiente propiedad Or_Φ, de *ser el código de una oración perteneciente a* Φ, es una propiedad recursivamente enumerable.

§ 5.29. Definición (Teorías recursivamente axiomatizables). Una teoría es *recursivamente axiomatizable* cuando existe un conjunto de axiomas para ella que constituye un conjunto r.e. de oraciones.

Naturalmente, siempre existe un conjunto de axiomas para cualquier teoría Δ, ya que por definición de teoría, $teo\,(\Delta) = \Delta$ (cf. § 2.84, p. 74).

El mérito de las teorías recursivamente axiomatizables es que poseen algún conjunto de axiomas que es recursivamente enumerable, es decir, cuyos códigos se pueden generar mecánicamente por medio de un conjunto finito de instrucciones.

§ 5.30. Lema (Teorías recursivamente axiomatizables). *Si Δ es una teoría recursivamente axiomatizable, entonces existe un conjunto recursivo de axiomas para Δ.*

Prueba. Sea Δ una teoría recursivamente axiomatizable, y sea Γ un conjunto r.e. de axiomas para Δ. Si Γ es finito, entonces Or_Γ es obviamente una propiedad recursiva: en efecto, para saber si un número natural dado es el código de una oración de Γ, bastaría con compararlo con cada uno de los códigos de las oraciones de Γ, uno detrás de otro, hasta terminar. En este caso, por tanto, el resultado queda probado.

Supongamos ahora que Γ es infinito. Como se trata de un conjunto r.e. de oraciones, podemos tomar un procedimiento de enumeración para el mismo, y considerar el orden en el que van apareciendo los códigos de las oraciones que este conjunto contiene:

$$\underline{\sharp\,\alpha_1}\,,\quad \sharp\,\underline{\alpha_2}\,,\quad \sharp\,\underline{\alpha_3}\,,\quad \ldots$$

A continuación construimos un nuevo conjunto de oraciones, Φ, componiendo conjunciones cada vez más largas, a partir del orden en que vienen dadas las oraciones de Γ en esa enumeración. En concreto, poniendo simplemente:

$$\Phi \quad = \quad \{\,\underline{\alpha_1}\,,\, \underline{\alpha_1 \wedge \alpha_2}\,,\, \underline{\alpha_1 \wedge \alpha_2 \wedge \alpha_3}\,,\, \ldots\}$$

Es evidente que las consecuencias lógicas de Γ y Φ son exactamente las mismas, por lo que tenemos $teo\,(\Gamma) = teo\,(\Phi)$. Por lo tanto, tenemos también $teo\,(\Phi) = \Delta$, con lo que Φ es claramente un conjunto de axiomas para Δ.

Pero además, dado que las oraciones de Φ aparecen en orden de longitud estrictamente creciente, sus correspondientes códigos serán también números crecientes, de menor a mayor.

Por consiguiente, para determinar si un número natural es el código de una oración de Φ o no, basta con ir comparándolo con los códigos de las oraciones que se vayan obteniendo en la enumeración, hasta encontrar la primera que sea mayor que el número dado. Si para entonces no ha aparecido, podemos concluir que dicho número no es el código de ninguna oración de Φ, y no hay necesidad de buscar más. Lo cual prueba que Or_Φ es una propiedad recursiva.

§ 5.31. Definición (Códigos de secuencias). A cada secuencia finita de fórmulas de $\mathscr{A}$ le asignamos también un *código* o *número de Gödel*, simplemente concatenando los códigos de las fórmulas que la componen, dados en su representación decimal habitual, y por el orden en que aparecen en ella.

§ 5.32. Definición (La relación *Ded*). Sea Φ un conjunto de oraciones de $\mathscr{A}$. Entonces dos números naturales a y b cumplen la relación Ded_Φ, exactamente en aquellos casos en los que a sea el código de una oración, y b sea el código de una secuencia de fórmulas que constituya una deducción de dicha oración desde el conjunto Φ.

§ 5.33. Práctica* (La relación *Ded*).

(a) Indicar cómo se podría definir un procedimiento mecánico para identificar, dado el código de una secuencia de fórmulas, las distintas fórmulas que componen ésta.

(b) Mostrar que si Φ es un conjunto recursivo de oraciones, entonces la relación Ded_Φ es recursiva.

§ 5.34. Teorema (Teorías recursivamente axiomatizables). *Cualquier teoría recursivamente axiomatizable es un conjunto de oraciones recursivamente enumerable.*

Prueba. Si Δ es una teoría recursivamente axiomatizable, entonces por el *Lema* § 5.30 existirá un conjunto recursivo de axiomas, Φ, para Δ. Por lo tanto, aplicando la Práctica precedente, vemos que Ded_Φ es una relación recursiva, y por consiguiente r.e.

En consecuencia, existe un procedimiento de enumeración para todos los pares de números naturales a, b tales que a es el código de una oración y b es el código de una deducción de esa oración desde Φ. Si ahora tomamos de esa enumeración sólo el primer componente de cada par que se vaya generando, y omitimos las repeticiones que puedan ir ocurriendo, habremos obtenido una enumeración de los códigos de todas las oraciones deducibles de Φ en el cálculo de predicados de $\mathscr{A}$. Es decir: una enumeración de todas las oraciones que pertenecen a Δ, puesto que Φ es un conjunto de axiomas para Δ.

Y con ello hemos mostrado que Or_Δ es una propiedad r.e., y por tanto la propia Δ es también r.e.

§ 5.35. Observación (Teorías recursivamente axiomatizables). En resumen, si una teoría es recursivamente axiomatizable, entonces no sólo posee un conjunto de axiomas recursivamente enumerable, sino que por el *Lema* § 5.30 posee algún otro que es recursivo, y por el teorema precedente, además, *toda* la teoría es un conjunto recursivamente enumerable de oraciones.

Por tanto, si una teoría Δ es recursivamente axiomatizable, entonces existirá un conjunto finito de instrucciones precisas para generar, uno a uno, todos los números naturales que cumplen Or_Δ, es decir, que son códigos de alguna oración perteneciente a Δ. Y a partir de cada uno de estos números es inmediato obtener la oración de la cual cada uno es código.

§ 5.36. Teorema (Ω no es recursivamente axiomatizable). *La teoría Ω no es recursivamente axiomatizable.*

Prueba. Por el teorema de Tarski, Or_Ω no es aritmética. Por tanto, por el *Corolario* § 5.13 (p. 100) no es r.e., y por el *Teorema* § 5.34 no puede ser recursivamente axiomatizable.

§ 5.37. Observación (Ω no es recursivamente axiomatizable). Nuestra introducción de la teoría Ω (en § 3.14, p. 79) se basó en una referencia explícita al modelo $\mathscr{N}$ de los números naturales, y en este sentido, supuso cierto compromiso con la existencia de esa estructura.

Sin embargo, el teorema precedente, que es una consecuencia directa del teorema de Tarski, nos hace ver ahora que de hecho no es posible caracterizar la teoría Ω de una manera puramente mecánica, o *sintáctica*, es decir, que no es posible caracterizar Ω sin hacer referencia a alguna interpretación suya.

Recordemos que ya habíamos renunciado a caracterizar unívocamente nuestro modelo sobreentendido $\mathscr{N}$ (§ 3.27, p. 85). Inclusive la teoría Ω, que recogía *toda* la verdad sobre $\mathscr{N}$ expresable en el lenguaje $\mathscr{A}$, tenía modelos no estándar.

Pero lo que descubrimos ahora es que ni siquiera podemos dar una especificación mecánica *de la propia teoría* Ω. Es decir: la teoría Ω, como conjunto o selección de oraciones de $\mathscr{A}$, representa una abstracción tan grande como el propio modelo $\mathscr{N}$: no hay forma de decir qué oraciones pertenecen a Ω sin mencionar para ello a $\mathscr{N}$, esto es, sin utilizar para ello la estructura abstracta de los números naturales y sus propiedades.

La cuestión que se plantea a continuación, claro está, es qué parte de la teoría Ω sí es recursivamente axiomatizable. Es decir, hasta qué fragmento parcial de Ω se puede axiomatizar, esto es, someter al tratamiento formal. Y ésa será, en efecto, una de nuestras principales preocupaciones a lo largo del Módulo siguiente.

§ 5.38. Teorema (Representación débil). *Cualquier relación débilmente representable en una teoría recursivamente axiomatizable es una relación r.e.*

Prueba. Sean Δ una teoría de $\mathscr{A}$ recursivamente axiomatizable, R una relación n-aria, y $\underline{\alpha} \in \Upsilon_n$ una fórmula que representa débilmente a R en Δ. Por el *Teorema* § 5.34 Δ es un conjunto r.e. de oraciones, y por tanto la propiedad Or_Δ también será r.e.

Entonces ha de existir un procedimiento de enumeración para Or_Δ, es decir, un procedimiento para generar una a una todas las oraciones de Δ.

Ahora bien, para cada una de estas oraciones, es inmediato determinar si son de la forma $\underline{\alpha[s_{\vec{a}}]}$ para alguna secuencia $\vec{a} \in \mathbb{N}^n$ o no, simplemente descodificándola y comprobándolo. Y en el caso en que esto suceda, seleccionar la secuencia $\vec{a}$ correspondiente.

Es obvio que el resultado, borrando repeticiones según ocurran, será una enumeración de todas las secuencias $\vec{a} \in \mathbb{N}^n$ tales que la correspondiente oración $\underline{\alpha[s_{\vec{a}}]}$ pertenece a Δ. En otras palabras, puesto que $\underline{\alpha}$ representa débilmente a R en Δ: una enumeración de todas las secuencias $\vec{a} \in \mathbb{N}^n$ que cumplen $R(\vec{a})$.

Y con ello hemos demostrado que R es r.e.

§ 5.39. Práctica* (Representación fuerte). Sea Δ una teoría de $\mathscr{A}$ recursivamente axiomatizable y consistente, y sea R una relación n-aria.

(a) Probar que si una fórmula $\underline{\alpha}$ representa fuertemente a R en Δ, entonces la fórmula $\underline{\neg\alpha}$ representará fuertemente a la relación complementaria de R.

(b) Probar que si R es fuertemente representable en Δ, entonces es una relación recursiva.

MÓDULO 6

El teorema de Church y los teoremas de Gödel

La aritmética de Peano de primer orden

§ 6.1. Definición (La teoría Π_0). Para cualesquiera números naturales n y m, son *axiomas de la teoría* Π_0 todas las oraciones de $\mathscr{A}$ que tengan alguna de estas cuatro formas:

$$\underline{s_n + s_0 = s_n} \qquad \underline{s_n + s_{m+1} = s\,(s_n + s_m)}$$

$$\underline{s_n \times s_0 = s_0} \qquad \underline{s_n \times s_{m+1} = s_n + (s_n \times s_m)}$$

para cualesquiera números naturales n y m.

Entonces, la *teoría* Π_0 es aquella teoría generada por el conjunto de axiomas que se acaba de describir. Es decir: Π_0 es el conjunto de todas las oraciones de $\mathscr{A}$ que son deducibles de esos axiomas en el cálculo de predicados.

§ 6.2. Observación (La teoría Π_0). Cualquiera que sepa sumar y multiplicar apreciará que todos los axiomas de Π_0 son verdaderos en el modelo estándar $\mathscr{N}$. En efecto, para cualesquiera $n, m \in \mathbb{N}$, tenemos:

$$n + 0 = n \qquad n + (m+1) = s\,(n+m)$$

$$n \cdot 0 = 0 \qquad n \cdot (m+1) = n + (n \cdot m)$$

Por lo tanto, *todas* las oraciones de Π_0 serán verdaderas en el modelo $\mathscr{N}$, es decir: Π_0 es lo que definimos en § 3.15 (p. 79) como una teoría *correcta con respecto a* $\mathscr{N}$.

Por otra parte, aunque los axiomas de Π_0 son obviamente infinitos, es claro que dado un número natural cualquiera, podemos verificar mecánicamente si es o no el código de un axioma de Π_0, viendo primero si es el código de una oración, y comparándolo en tal caso con cada uno de los cuatro grupos de la definición precedente. Por lo tanto, este conjunto de axiomas es recursivo, y por consiguiente r.e., de modo que Π_0 es una teoría recursivamente axiomatizable.

§ 6.3. Práctica* (Las tablas formales en Π_0). Probar que Π_0 contiene las contrapartidas formales de las tablas de la suma y la multiplicación. Es decir, que para cualesquiera $n, m \in \mathbb{N}$, la teoría Π_0 contiene todas las oraciones de la forma

(a) $\underline{s_n + s_m = s_{n+m}}$ (b) $\underline{s_n \times s_m = s_{n \cdot m}}$

Para ello, efectuar cada prueba separadamente, por inducción débil sobre m.

§ 6.4. Lema (Atómicas verdaderas en Π_0). Si $\underline{t}$ es un término cerrado y la oración atómica $\underline{t = s_n}$ es verdadera en $\mathscr{N}$, entonces dicha oración pertenece a Π_0.

Prueba. Procedemos por inducción fuerte sobre el grado de $\underline{t}$, tratando separadamente las distintas formas que puede tomar $\underline{t}$ de acuerdo con la *Observación* § 3.3 (p. 76), excepto el caso de que sea una variable sola, lo cual es imposible al tratarse de un término cerrado.

En primer lugar supongamos que $\underline{t}$ sea la constante $\underline{0}$. La única oración verdadera en este caso es $\underline{0 = s_0}$, esto es, la oración $\underline{0 = 0}$, que es una verdad lógica. Por lo tanto dicha oración es una consecuencia trivial de cualquier conjunto de oraciones, y en particular de Π_0. Y al ser Π_0 una teoría, tendrá que pertenecer a ella.

En segundo lugar supongamos que $\underline{t}$ sea de la forma $\underline{sr}$, para algún otro término cerrado $\underline{r}$ de $\mathscr{A}$. Como partimos del supuesto de que la oración $\underline{sr = s_n}$ es verdadera en $\mathscr{N}$, eso significa que forzosamente $n \geqslant 1$. Por lo tanto podemos considerar a su vez el término $\underline{s_{n-1}}$, y la correspondiente oración $\underline{r = s_{n-1}}$, que será lógicamente equivalente a la anterior.

Entonces, al ser $\underline{sr = s_n}$ verdadera en $\mathscr{N}$, también lo será $\underline{r = s_{n-1}}$. Aplicando la hipótesis de inducción tenemos $\underline{r = s_{n-1}} \in \Pi_0$, y al ser Π_0 una teoría deberá tener a su vez a la equivalente a ésta, es decir: $\underline{sr = s_n} \in \Pi_0$.

En tercer lugar supongamos que $\underline{t}$ sea de la forma $\underline{r + r'}$, donde $\underline{r}$ y $\underline{r'}$ son ambos términos cerrados. Tomemos como valores respectivos de estos términos bajo $\mathscr{N}$ los naturales n, m y m', es decir:

$$\underline{t}^{\mathscr{N}} = n \qquad \underline{r}^{\mathscr{N}} = m \qquad \underline{r'}^{\mathscr{N}} = m'$$

donde naturalmente $n = m + m'$.

Entonces las oraciones

$$\underline{r = s_m} \qquad \underline{r' = s_{m'}}$$

serán ambas verdaderas en $\mathscr{N}$, y por consiguiente, por hipótesis de inducción, pertenecerán a Π_0.

Ahora bien: una consecuencia lógica inmediata de esas dos oraciones es

$$\underline{r + r' = s_m + s_{m'}}$$

que también habrá de pertenecer a Π_0, al ser una teoría. Y por otro lado, como Π_0 contiene la oración

$$\underline{s_m + s_{m'} = s_n}$$

al ser verdadera (por la *Práctica* § 6.3), podemos extraer como consecuencia lógica inmediata la oración $\underline{r + r' = s_n}$ que, una vez más, también debe estar en Π_0, al ser ésta una teoría.

Y en cuarto y último lugar, en el caso en que $\underline{t}$ sea de la forma $\underline{r \times r'}$, se procede de forma enteramente análoga al caso anterior.

§ 6.5. Teorema (Atómicas verdaderas en Π_0). *Cualquier oración atómica verdadera en $\mathscr{N}$ pertenece a Π_0.*

Prueba. Sean $\underline{t}$ y $\underline{r}$ dos términos cerrados, y supongamos que la oración $\underline{t = r}$ es verdadera en $\mathscr{N}$, es decir, que $\underline{t}^{\mathscr{N}} = \underline{r}^{\mathscr{N}}$. Sea $n \in \mathbb{N}$ el valor de $\underline{t}^{\mathscr{N}}$, y por tanto también el de $\underline{r}^{\mathscr{N}}$.

Por el lema precedente tenemos que las oraciones $\underline{t = s_n}$ y $\underline{r = s_n}$ pertenecen ambas a Π_0. Y una consecuencia lógica inmediata de ellas es $\underline{t = r}$, que también habrá de pertenecer por tanto a Π_0.

§ 6.6. Lema (Existenciales verdaderas en Π_0). *Cualquier oración verdadera en $\mathscr{N}$ y que tenga la forma*

$$\underline{\exists x_1 x_2 \ldots x_n \, \beta}$$

donde $\underline{x_1}$, $\underline{x_2}$, ..., $\underline{x_n}$ son variables distintas, y $\underline{\beta}$ es una fórmula atómica de $\mathscr{A}$, pertenece a Π_0.

Prueba. Procedemos por inducción fuerte sobre el número n de cuantificadores. Si $n = 0$, entonces la oración es de la forma $\underline{\beta}$ sin más, y el resultado se sigue de inmediato del teorema precedente. Suponiendo entonces que $n \geqslant 1$, consideremos la fórmula resultante de suprimir el primer cuantificador:

$$\underline{\alpha} \quad = \quad \underline{\exists x_2 \ldots x_n \, \beta}$$

De tal forma que

$$\underline{\exists x_1 x_2 \ldots x_n \, \beta} \quad = \quad \underline{\exists x_1 \alpha}$$

Al ser $\underline{\exists x_1 \alpha}$ una oración, es claro que la única variable que puede ocurrir libre en $\underline{\alpha}$ es $\underline{x_1}$. Por tanto $\underline{\alpha \, [x_1/s_m]}$ será a su vez una oración, para cualquier numeral $\underline{s_m}$.

Ahora bien: si $\underline{\exists x_1 \alpha}$ es verdadera en $\mathscr{N}$, deberá existir algún $m \in \mathbb{N}$ en concreto, tal que la oración $\underline{\alpha \, [x_1/s_m]}$ sea también verdadera en $\mathscr{N}$. Y como esta última oración tiene sólo $n - 1$ cuantificadores, por hipótesis de inducción pertenecerá a Π_0.

Por último, dado que la oración $\underline{\exists x_1 \alpha}$ es una consecuencia lógica de $\underline{\alpha \, [x_1/s_m]}$, también deberá pertenecer a Π_0, por el hecho de ser Π_0 una teoría.

§ 6.7. Teorema (Representación débil en Π_0). *Cualquier relación r.e. es débilmente representable en Π_0.*

Prueba. Sea Q una relación n-aria y r.e. Por el teorema de Matiyasevich, Q será una relación diofántica, y por tanto existirá una relación elemental R tal que para cualquier $\vec{a} \in \mathbb{N}^n$,

$$Q(\vec{a}) \qquad \text{si y sólo si} \qquad \text{existe una secuencia } \vec{b} \in \mathbb{N}^m \text{ tal que } R(\vec{a}, \vec{b})$$

De acuerdo con la prueba del *Lema* § 5.10 (p. 99), al ser R elemental, existe una fórmula atómica $\underline{\beta}$ que la representa en Ω. Y por tanto, de acuerdo ahora con la prueba del *Teorema* § 5.12 (p. 100), la fórmula

$$\underline{\alpha} \quad = \quad \underline{\exists \, v_{n+1} \, v_{n+2} \, \ldots \, v_{n+m} \;\; \beta}$$

representa a Q en Ω. Es decir, que para cualquier $\vec{a} \in \mathbb{N}^n$,

$$Q(\vec{a}) \qquad \text{si y sólo si} \qquad \underline{\alpha \, [s_{\vec{a}}]} \in \Omega$$

Ahora supongamos que, en efecto, sucede $Q(\vec{a})$. Entonces la oración $\underline{\alpha\,[s_{\vec{a}}]}$ pertenecerá a Ω, por tanto será verdadera en $\mathscr{N}$, y puesto que se ajusta a las condiciones del lema precedente, pertenecerá a Π_0.

Y por otra parte, si $\sim Q(\vec{a})$, entonces $\underline{\alpha\,[s_{\vec{a}}]} \notin \Omega$, y por tanto la oración $\underline{\alpha\,[s_{\vec{a}}]}$ será falsa en $\mathscr{N}$. Pero entonces no podrá pertenecer tampoco a Π_0, por ser ésta una teoría correcta con respecto a $\mathscr{N}$ (*Observación* § 6.2).

Y ello prueba que $\underline{\alpha}$ representa débilmente a Q en Π_0.

§ 6.8. Corolario (Representación débil en Π_0). *Cualquier relación r.e. es débilmente representable en cualquier teoría de $\mathscr{A}$ que sea correcta con respecto a $\mathscr{N}$ y que incluya a Π_0.*

Prueba. Sea Q una relación n-aria r.e., y sea Δ cualquier teoría de $\mathscr{A}$ que sea correcta con respecto a $\mathscr{N}$ y que incluya a Π_0. Por el teorema precedente, existe una fórmula $\underline{\alpha}$ que representa débilmente a Q en Π_0, y, según se ha visto en la prueba de este teorema, también en Ω.

Entonces, para cualquier $\vec{a} \in \mathbb{N}^n$, si sucede $Q(\vec{a})$ tendremos $\underline{\alpha\,[s_{\vec{a}}]} \in \Pi_0$, y por lo tanto, dado que Π_0 está incluida en Δ, también $\underline{\alpha\,[s_{\vec{a}}]} \in \Delta$.

Y si sucede $\sim Q(\vec{a})$, entonces tendremos $\underline{\alpha\,[s_{\vec{a}}]} \notin \Omega$, es decir, $\underline{\alpha\,[s_{\vec{a}}]}$ será una oración falsa en $\mathscr{N}$, y por lo tanto no podrá pertenecer a Δ, al ser ésta una teoría correcta con respecto a $\mathscr{N}$.

De tal manera que la misma fórmula $\underline{\alpha}$ también representa débilmente a Q en Δ.

§ 6.9. Observación (Problema 9). La teoría Π_0 no contiene todas las oraciones verdaderas en $\mathscr{N}$. Un ejemplo muy sencillo de obtener nos lo muestra el *Problema 9* del listado de *Problemas propuestos* (p. 13). Ahora es el momento apropiado para plantearse la respuesta a dicho problema.

§ 6.10. Convención (Abreviatura especial). Si $\underline{r}$ y $\underline{t}$ son términos cualesquiera de $\mathscr{A}$, y $\underline{x}$ es la primera variable, en el orden de las variables de $\mathscr{A}$, que no aparece en ninguno de los dos, entonces:

- La fórmula $\underline{r \geqslant t}$ será por definición igual a la fórmula $\underline{\exists x\,(\,r = t + x\,)}$.

Según esta convención, la fórmula $\underline{v_1 \geqslant v_2}$, por ejemplo, no es otra que

$$\underline{\exists v_3\,(\,v_1 = v_2 + v_3\,)}$$

la cual, según vimos en la *Práctica* § 5.9(a) (p. 98), representa en Ω la relación de *ser mayor o igual que.*

§ 6.11. Definición (La teoría Π_1). Son *axiomas de la teoría* Π_1 todos los axiomas de la teoría Π_0, más todas las oraciones de $\mathscr{A}$ que tengan alguna de las siguientes tres formas:

$$\underline{\forall v_1\,(\,s_n \geqslant v_1 \;\rightarrow\; (s_0 = v_1 \vee s_1 = v_1 \vee \ldots \vee s_n = v_1)\,)} \qquad \underline{\forall v_1\,(s_n \geqslant v_1 \vee v_1 \geqslant s_n)}$$

$$\text{y} \quad \underline{s_n \neq s_m} \quad \text{para cualesquiera números naturales } \textit{distintos } n \text{ y } m$$

Entonces, la *teoría* Π_1 es aquella teoría generada por el conjunto de axiomas que se acaba de definir. Es decir: Π_1 es el conjunto de todas las oraciones de $\mathscr{A}$ que son deducibles de esos axiomas en el cálculo de predicados.

§ 6.12. Observación (La teoría Π_1). Al igual que sucedía con su predecesora Π_0, es evidente que la teoría Π_1 constituye también una teoría correcta con respecto a $\mathscr{N}$, y por supuesto, que su conjunto de axiomas es recursivo, con lo que resulta también una teoría recursivamente axiomatizable.

§ 6.13. Lema (Representación fuerte en Π_1). *Si R y R' son relaciones n-arias y r.e., entonces existe una fórmula $\underline{\alpha} \in \Upsilon_n$ tal que para todo $\vec{a} \in \mathbb{N}^n$,*

$$\textit{si sucede} \quad R(\vec{a}) \quad y \quad \sim R'(\vec{a}), \quad \textit{entonces} \quad \underline{\alpha\,[s_{\vec{a}}]} \in \Pi_1$$

$$\textit{si sucede} \quad \sim R(\vec{a}) \quad y \quad R'(\vec{a}), \quad \textit{entonces} \quad \underline{\neg\alpha\,[s_{\vec{a}}]} \in \Pi_1$$

Prueba. La prueba de este importante lema es larga y tediosa, por lo que la omitimos.

§ 6.14. Teorema (Representación fuerte en Π_1). *Cualquier relación recursiva es fuertemente representable en Π_1.*

Prueba. Sea R una relación recursiva n-aria, y tomemos como R' a la relación complementaria de R. Es inmediato entonces que tanto R como R' son n-arias, y también r.e., según vimos en la *Práctica* § 4.10 (p. 92).

Por tanto basta con aplicar el lema precedente, y es inmediato comprobar que la fórmula $\underline{\alpha}$ allí mencionada representará fuertemente a R en Π_1.

§ 6.15. Corolario (Representación fuerte en Π_1). *Cualquier relación recursiva es fuertemente representable en cualquier teoría de $\mathscr{A}$ que incluya a Π_1.*

Prueba. Inmediata a partir del teorema precedente.

§ 6.16. Definición (La teoría Π_2). Son *axiomas de la teoría* Π_2 las siguientes nueve oraciones de $\mathscr{A}$:

$$\underline{\forall v_1\,(\,v_1 + s_0 = v_1\,)}$$

$$\underline{\forall v_1 v_2\,(\,v_1 + s v_2 = s(v_1 + v_2)\,)}$$

$$\underline{\forall v_1\,(\,v_1 \times s_0 = s_0\,)}$$

$$\underline{\forall v_1 v_2\,(\,v_1 \times s v_2 = v_1 + (v_1 \times v_2)\,)}$$

$$\underline{\forall v_1\,(\,s v_1 \neq s_0\,)}$$

$$\underline{\forall v_1 v_2\,(\,s v_1 = s v_2 \;\rightarrow\; v_1 = v_2\,)}$$

$$\underline{\forall v_1\,(\,s_0 \geqslant v_1 \;\rightarrow\; s_0 = v_1\,)}$$

$$\underline{\forall v_1 v_2\,(\,s v_2 \geqslant v_1 \;\rightarrow\; (v_2 \geqslant v_1 \vee s v_2 = v_1)\,)}$$

$$\underline{\forall v_1 v_2\,(\,v_1 \geqslant v_2 \vee v_2 \geqslant v_1\,)}$$

Entonces, la *teoría* Π_2 es aquella teoría generada por dichos nueve axiomas.

§ 6.17. Definición (La aritmética de Peano). Son *axiomas de la aritmética de Peano de primer orden* los nueve axiomas de la teoría Π_2, más todas las oraciones de $\mathscr{A}$ que tienen la forma que a continuación se indica, y que se denominan *"axiomas de inducción"*:

$$\underline{\forall v_2 v_3 \ldots v_n \left((\alpha\,[s_0] \,\wedge\, \forall v_1(\alpha \to \alpha\,[v_1/sv_1])) \;\to\; \forall v_1 \alpha \right)}$$

donde n es cualquier número natural $\geqslant 1$ y $\underline{\alpha}$ cualquier fórmula tal que $\underline{\alpha} \in \Upsilon_n$.

Entonces, la *aritmética de Peano de primer orden* (abreviadamente, "Π" a secas), es aquella teoría generada por todos esos axiomas.

§ 6.18. Observación (Los axiomas de inducción). Para el caso $n = 1$, los axiomas de inducción adoptan la forma

$$\underline{(\,\alpha\,[s_0] \,\wedge\, \forall v_1(\alpha \to \alpha\,[v_1/sv_1])\,) \;\to\; \forall v_1 \alpha}$$

donde $\underline{\alpha} \in \Upsilon_1$, y por tanto $\underline{\alpha}$ representa una propiedad en Ω. Es evidente entonces que el axioma viene a expresar precisamente el hecho de que esa propiedad cumple con el principio de inducción débil, según se describió en § 1.50 (p. 40).

§ 6.19. Observación (Las teorías Π_2 y Π). Es una cuestión rutinaria y sencilla el comprobar que todos los axiomas de Π_1 son deducibles en la teoría Π_2. Así pues Π_1 está incluida en Π_2, y por lo tanto en Π.

Tampoco es difícil comprobar que tanto Π_2 como Π son teorías correctas con respecto a $\mathscr{N}$. Y desde luego, está a la vista que ambas son recursivamente axiomatizables.

De ello se sigue de inmediato que ninguna de estas teorías abarca todas las oraciones de Ω, es decir, todas las oraciones verdaderas en $\mathscr{N}$ (*Teorema* § 5.36, p. 108). Sin embargo, a diferencia de sus predecesoras, Π en particular es una teoría verdaderamente rica, y requiere una gran dosis de ingenio el encontrar oraciones verdaderas en $\mathscr{N}$ que no pertenezcan a ella.

§ 6.20. Observación (La aritmética de Peano). La teoría Π debe su nombre al matemático italiano Giuseppe Peano, que fue el primero en presentar, en 1889, una axiomatización de las propiedades de los números naturales en un lenguaje completamente formalizado. Tal axiomatización, compuesta de cinco axiomas que se conocen generalmente como *"axiomas de Peano"*, se basó a su vez en una anterior, propuesta por el matemático alemán Richard Dedekind un año atrás. Por lo demás, cualquiera de las dos difiere mucho en forma y alcance de nuestra teoría Π tal y como aquí la hemos presentado.

El teorema de Church

§ 6.21. Definición (Teorías decidibles). Una teoría es *decidible* cuando forma un conjunto recursivo de oraciones. Y es *indecidible* en caso contrario.

§ 6.22. Lema (Teorías indecidibles). *Si Δ es una teoría de $\mathscr{A}$ en la cual cualquier propiedad recursiva es débilmente representable, entonces Δ es indecidible.*

Prueba. La prueba de este lema tiene una estructura enteramente similar a la del teorema de Tarski, por lo que la dejamos como Práctica, que a continuación se propone.

§ 6.23. Práctica* (Teorías indecidibles). Probar el lema que se acaba de enunciar.

§ 6.24. Corolario (Indecidibilidad de Π_1). *Cualquier teoría de $\mathscr{A}$ que sea consistente y que incluya a Π_1 es indecidible.*

Prueba. Sea Δ una teoría de $\mathscr{A}$ que sea consistente e incluya a Π_1. Por el *Corolario* § 6.15 cualquier relación recursiva será fuertemente representable en Δ. Al ser Δ consistente, la representación fuerte implica la débil (*Observación* § 5.4, p. 98). Entonces el resultado se sigue de inmediato del lema precedente.

§ 6.25. Teorema (Indecidibilidad de Π_2). *Cualquier teoría de $\mathscr{A}$ cuya unión con Π_2 sea consistente, es indecidible.*

Prueba. Sea Δ cualquier teoría de $\mathscr{A}$ cuya unión con Π_2 forme un conjunto consistente de oraciones. Sea ahora Φ la teoría de ese conjunto, es decir:

$$\Phi = teo\,(\Delta \cup \Pi_2)$$

Naturalmente, Φ incluye a Π_2, y por lo tanto a Π_1, por lo que Φ es claramente indecidible, según el corolario precedente.

Sea ahora $\underline{\pi}$ la oración formada por la conjunción de los nueve axiomas de Π_2. Entonces es claro que una oración cualquiera $\underline{\alpha}$ pertenecerá a Φ si y sólo si es deducible de Δ y $\underline{\pi}$, con lo que a su vez:

$$\begin{array}{llllll}
\underline{\alpha} \in \Phi & \text{si y sólo si} & \Delta \cup \{\underline{\pi}\} & \vdash & \underline{\alpha} & \\
 & \quad '' & \Delta \cup \{\underline{\pi}\} & \models & \underline{\alpha} & \text{(por el } \textit{Teorema}\ \S\, 2.70\text{, p. 71)} \\
 & \quad '' & \Delta & \models & \underline{\pi \to \alpha} & \text{(obvio por } \S\, 2.35\text{, p. 58)} \\
 & \quad '' & & & \underline{\pi \to \alpha} \in \Delta & \text{(al ser } \Delta \text{ una teoría)}
\end{array}$$

En otras palabras: una oración $\underline{\alpha}$ pertenecerá a Φ si y sólo si la correspondiente oración $\underline{\pi \to \alpha}$ pertenece a Δ. Por lo tanto, si Δ fuese recursiva también lo sería Φ, cosa que es imposible al ser Φ una teoría indecidible.

Así pues Δ es también una teoría indecidible.

§ 6.26. Teorema (Teorema de Church). *La teoría Λ es indecidible.*

Prueba. $\Lambda \cup \Pi_2$ es la propia teoría Π_2, que es correcta con respecto a $\mathscr{N}$, y por lo tanto consistente.

§ 6.27. Observación (El teorema de Church). El teorema precedente fue obtenido por Church en 1936, y se trata de un resultado que versa exclusivamente sobre la lógica de primer orden, aunque para establecerlo, nosotros nos hayamos apoyado en la exploración del lenguaje $\mathscr{A}$ y sus interpretaciones.

La teoría Λ es, como vimos en su momento, el conjunto de todas las oraciones de $\mathscr{A}$ que son deducibles del conjunto vacío de premisas (*Definición* § 3.12, p. 79). Por lo tanto es trivialmente una teoría recursivamente axiomatizable, y por consiguiente un conjunto r.e. de oraciones (*Teorema* § 5.34, p. 107).

Ahora bien, lo que nos dice el teorema de Church es que *no* es un conjunto recursivo de oraciones. Es decir: que no existe un procedimiento de decisión para determinar de un modo mecánico si un número natural dado es o no el código de una oración de Λ.

Ello implica, en definitiva, que no puede existir un procedimiento mecánico para determinar si una oración de $\mathscr{A}$ es o no deducible del conjunto vacío de premisas, o lo que es lo mismo, que no puede existir un procedimiento mecánico para determinar si una oración de $\mathscr{A}$ es o no una verdad lógica.

Al ser Λ r.e., lo que sí podemos conseguir es un procedimiento de enumeración para generar, una por una, todas las oraciones de Λ, es decir, todas las verdades lógicas. Por lo tanto, si una oración concreta por la que nos preguntamos, resulta ser de hecho una verdad lógica, entonces al poner en marcha la enumeración, dicha oración tendrá que aparecer en ella, tarde o temprano. Pero si la oración en cuestión resulta que *no* es una verdad lógica, entonces la enumeración proseguirá hasta el infinito, la oración no aparecerá, y nosotros nos quedaremos sin saber si es que la oración va a aparecer más tarde, o es que no va a aparecer nunca. Esto es: el procedimiento nos dejará sin respuesta.

Claro que si la oración en cuestión no es excesivamente larga o complicada, siempre podremos, utilizando el ingenio, encontrar alguna forma de establecer definitivamente si es una verdad lógica o no. Por ejemplo, buscando posibles modelos, contramodelos, etc. Pero esto es ya algo que hacemos de forma particular, caso por caso, sin atenernos a un procedimiento general. Porque un procedimiento general, como vemos, no puede haberlo.

§ 6.28. Observación (Indecidibilidad de la lógica de primer orden). En definitiva, no puede existir tampoco ningún procedimiento general para decidir cuáles son las verdades lógicas de cualquier lenguaje de primer orden, ni, por supuesto, para decidir si una fórmula dada es o no una consecuencia lógica de un conjunto de fórmulas. Y esto es lo que se suele indicar condensadamente diciendo que *"la lógica de primer orden es indecidible"*.

Ahora bien: si la lógica de primer orden es la teoría del razonamiento deductivo correcto, ello significa a su vez que no puede existir un procedimiento mecánico para decidir de un modo general y en todos los casos, cuándo un razonamiento deductivo es correcto o incorrecto.

Se trata, como vemos, de un resultado de interés e importancia evidentes, en lo que supone de limitación fundamental a la hora de representar formalmente las condiciones de nuestro deducir y razonar.

Así, sabemos que resulta imposible, desde luego, diseñar un programa de ordenador que resuelva correctamente y en todos los casos el problema de si una fórmula es una verdad lógica, o es una consecuencia lógica de otra. Se encontrarán a lo sumo estrategias parciales, más o menos eficaces, que, sin llegar a cometer errores, resuelvan correctamente algunos casos, y dejen sin respuesta otros. Exactamente como ocurre con un procedimiento de enumeración para las oraciones de Λ, según hemos explicado antes.

A pesar de lo cual, sin embargo, no está nada claro que la limitación que estamos comentando nos haga necesariamente superiores a los hombres sobre los ordenadores. Y ello es así por la sencilla razón de que *tampoco nosotros* parecemos estar dotados de un mecanismo general que nos permita decidir en todos los casos si una fórmula es o no una verdad lógica: muy al contrario, a poco que la fórmula sea suficientemente larga o complicada, nuestra capacidad de analizarla o calcular con ella tiende a desaparecer por completo.

§ 6.29. Corolario (Teorías indecidibles). *Cualquier teoría de $\mathscr{A}$ que sea correcta con respecto a $\mathscr{N}$ es indecidible.*

Prueba. Si Δ es una teoría correcta, también lo será $\Delta \cup \Pi_2$, y por tanto consistente.

Se sigue en definitiva, que ninguna teoría que refleje mínimamente la interpretación estándar $\mathscr{N}$ puede ser decidible.

§ 6.30. Observación (Problema 10). Este es el momento apropiado para plantearse la respuesta al *Problema 10* del listado de *Problemas propuestos* (p. 13).

Los teoremas de Gödel

§ 6.31. Observación (Los teoremas de Gödel). En 1931 Kurt Gödel publicó en un memorable artículo sus dos *teoremas de incompletitud*, de gran notoriedad y trascendencia, y que van a servir como cierre a la presente asignatura. Aunque Gödel es autor de muchos otros teoremas importantes en lógica, por ejemplo, sin ir más lejos, nuestro § 2.71 (p. 72), lo cierto es que cuando se habla de *"los teoremas de Gödel"* sin más, se está haciendo referencia concretamente a los teoremas de incompletitud, que vamos a comentar ahora. Y cuando se habla a secas de "*el* teorema de Gödel" se suele hacer referencia al primero de ellos, aunque el segundo reviste también una extraordinaria importancia.

§ 6.32. Observación (El primer teorema de incompletitud). El primer teorema de incompletitud comprende en realidad dos versiones, dos teoremas distintos pero íntimamente relacionados. La diferencia entre ambos tiene que ver con la oposición entre nociones sintácticas y nociones semánticas, de la que hablamos en su momento (cf. § 5.19, p. 102, y § 5.27, p. 105).

De acuerdo con la *versión semántica* de este teorema, cualquier teoría que sea recursivamente axiomatizable y correcta con respecto a $\mathscr{N}$, es necesariamente *incompleta* con respecto a $\mathscr{N}$. Es decir: que existe alguna oración verdadera en $\mathscr{N}$ que no pertenece a dicha teoría. Esto no puede sorprendernos a nosotros, que ya conocemos el teorema de Tarski, porque la única teoría correcta y completa con respecto a $\mathscr{N}$ es Ω, que no es recursivamente axiomatizable (cf. § 5.36, p. 108).

En efecto, la prueba de Gödel de este resultado tiene, como veremos, ciertas afinidades con la del teorema de Tarski, aunque la oración paradójica que se elabora en esta otra ocasión sí existe, y de hecho se podría especificar completamente, aunque nosotros no vamos a llegar a tanto detalle. En cualquier caso, se trata de una manera de incidir, desde otro ángulo, en lo que ya observamos en su momento a propósito de aquel teorema: que no puede haber ni siquiera un procedimiento de enumeración para el conjunto de oraciones verdaderas en $\mathscr{N}$. Es decir: que resulta imposible, utilizando el método formal, aprehender completamente la estructura $\mathscr{N}$.

En cuanto a la *versión sintáctica*, más profunda y general, establece que para cualquier teoría consistente, recursivamente axiomatizable y que incluya a Π_1, existe una oración tal que ni ella ni su negación pertenecen a la teoría. Esto es: que existe una oración del lenguaje que la teoría deja sin determinar. Se trata por lo tanto de un resultado en el que no se menciona en ningún momento a la interpretación o modelo $\mathscr{N}$, y que por consiguiente podemos enunciar y probar sin comprometernos en absoluto con la existencia de dicha estructura.

Volviendo a la orientación formalista, que describimos someramente en § 3.27 (p. 86), es posible que desde dicha perspectiva pudiéramos contentarnos con un fragmento totalmente formalizable de la aritmética, como pudiera ser la teoría Π. Ello haría indemostrables, dentro de nuestra teoría, algunas oraciones verdaderas sobre la estructura $\mathscr{N}$, de acuerdo con la versión semántica del primer teorema de incompletitud. Pero desde el punto de vista de alguien que no quiere comprometerse con la existencia de dicha estructura, la noción misma de "oración verdadera en $\mathscr{N}$" debería contemplarse con suspicacia, con lo que se trata de una pérdida asumible.

Lo que consigue establecer la versión sintáctica de este teorema, sin embargo, es que simplemente con que la teoría en cuestión sea consistente, recursivamente axiomatizable e incluya a Π_1, habrá oraciones que no se puedan determinar dentro de la teoría, porque ni ellas ni sus negaciones pertenezcan a la teoría. Es decir: que habrá oraciones, en definitiva problemas matemáticos, que no se puedan "resolver" dentro de la teoría escogida.

En resumidas cuentas, que la total formalización de cualquier parte de las matemáticas mínimamente interesante, como la aritmética, nos obligaría como vemos a trabajar con teorías, por decirlo de alguna manera, "inacabadas": teorías en las que no todas las cuestiones o problemas planteables tengan una solución determinada, que es exactamente lo que le pasa a Π. Y ello constituye sin duda un importante obstáculo para la concepción formalista.

§ 6.33. Teorema (Versión semántica del primer teorema de incompletitud). *Para cualquier teoría de $\mathscr{A}$ que sea recursivamente axiomatizable y correcta con respecto a $\mathscr{N}$, existe alguna oración verdadera en $\mathscr{N}$ que no pertenece a dicha teoría.*

Prueba. Sea Δ cualquier teoría de $\mathscr{A}$ correcta con respecto a $\mathscr{N}$ y recursivamente axiomatizable, y sea P aquella propiedad numérica tal que para todo $a \in \mathbb{N}$,

$$P(a) \qquad \text{si y sólo si} \qquad Or_{\Delta}\,(\,diag\,(a)\,)$$

Es decir: que un número natural a cumple la propiedad P si y sólo si el número resultante de efectuar la operación $diag\,(a)$ cumple la propiedad Or_{Δ}.

Puesto que Δ es recursivamente axiomatizable, la propiedad Or_Δ ha de ser r.e., y por la *Práctica* § 5.23(c) (p. 104) se sigue de inmediato que también P es r.e., y por lo tanto aritmética.

En definitiva, existirá una fórmula $\underline{\alpha} \in \Upsilon_1$ que represente a P en Ω, de lo cual se sigue, en particular, que para cualquier $a \in \mathbb{N}$

$$\underline{\neg\alpha\,[s_a]} \in \Omega \qquad \text{si y sólo si} \qquad \sim P(a)$$

Naturalmente, esta fórmula $\underline{\neg\alpha}$ tendrá un código, $\sharp\,\underline{\neg\alpha}$, que será un número natural. Por lo tanto tenemos:

$$
\begin{array}{llll}
\underline{\neg\alpha\,[s_{\sharp\neg\alpha}]} \in \Omega & \text{si y sólo si} & \sim Or_\Delta\,(\,diag\,(\sharp\,\underline{\neg\alpha})\,) & \text{(puesto que } \underline{\alpha} \text{ representa a } P \text{ en } \Omega) \\
 & \quad ” & \sim Or_\Delta\,(\,\sharp\,\underline{\neg\alpha\,[s_{\sharp\neg\alpha}]}\,) & \text{(por la } \textit{Observación}\ \S\,5.21\text{, p. 103)} \\
 & \quad ” & \underline{\neg\alpha\,[s_{\sharp\neg\alpha}]} \notin \Delta & \text{(por def. de } Or_\Delta\text{, cf. } \S\,5.24\text{, p. 104)}
\end{array}
$$

En definitiva, acabamos de demostrar que

$$\underline{\neg\alpha\,[s_{\sharp\neg\alpha}]} \in \Omega \qquad \text{si y sólo si} \qquad \underline{\neg\alpha\,[s_{\sharp\neg\alpha}]} \notin \Delta$$

Ahora bien: no puede ocurrir que $\underline{\neg\alpha\,[s_{\sharp\neg\alpha}]}$ sea una oración falsa en $\mathscr{N}$, ya que ello implicaría $\underline{\neg\alpha\,[s_{\sharp\neg\alpha}]} \notin \Omega$, y por lo tanto $\underline{\neg\alpha\,[s_{\sharp\neg\alpha}]} \in \Delta$, lo cual es imposible por ser Δ una teoría correcta con respecto a $\mathscr{N}$.

Por lo tanto $\underline{\neg\alpha\,[s_{\sharp\neg\alpha}]}$ ha de ser una oración verdadera en $\mathscr{N}$, y ello a su vez implica $\underline{\neg\alpha\,[s_{\sharp\neg\alpha}]} \in \Omega$, y por consiguiente $\underline{\neg\alpha\,[s_{\sharp\neg\alpha}]} \notin \Delta$. Es decir: en la oración $\underline{\neg\alpha\,[s_{\sharp\neg\alpha}]}$ hemos encontrado un ejemplo de oración verdadera en $\mathscr{N}$ que *no* pertenece a Δ.

§ 6.34. Observación (La versión semántica del primer teorema de incompletitud). A diferencia de la prueba del teorema de Tarski, y de la *Práctica* § 6.23, la fórmula en cuestión aquí considerada sí existe. Y puesto que dicha fórmula, $\underline{\alpha}$, representa a P en Ω, tenemos que para cada número natural a, la correspondiente oración $\underline{\neg\alpha\,[s_a]}$ será verdadera, y por lo tanto pertenecerá a Ω, si y sólo si $\sim P(a)$, es decir, si y sólo si $diag\,(a)$ *no* es el código de una oración de la teoría Δ.

En otras palabras, la oración $\underline{\neg\alpha\,[s_a]}$ viene a decir, leída indirectamente a través del método de codificación y de la teoría Ω, que $diag\,(a)$ no es el código de una oración de Δ.

Entonces, tomando como número el propio $\sharp\,\underline{\neg\alpha}$, tenemos que la oración $\underline{\neg\alpha\,[s_{\sharp\neg\alpha}]}$ viene a decir que $diag\,(\sharp\,\underline{\neg\alpha})$ no es el código de una oración de Δ. Por lo tanto, aplicando la *Observación* § 5.21 (p. 103) tenemos que a su vez la oración $\underline{\neg\alpha\,[s_{\sharp\neg\alpha}]}$ viene a decir algo así como *“yo no pertenezco a Δ”*.

Y es por ello que, siendo Δ una teoría correcta, dicha oración no puede “mentir”. Ya que si $\underline{\neg\alpha\,[s_{\sharp\neg\alpha}]}$ perteneciese a Δ entonces sería una oración falsa en $\mathscr{N}$, es decir, tendríamos $\underline{\neg\alpha\,[s_{\sharp\neg\alpha}]} \notin \Omega$, y ello haría de Δ una teoría incorrecta con respecto a $\mathscr{N}$. Por lo tanto se trata de una oración verdadera, y precisamente por esto no puede pertenecer a Δ.

§ 6.35. Teorema (Versión sintáctica del primer teorema de incompletitud). *Si Δ es una teoría de $\mathscr{A}$ que es consistente, recursivamente axiomatizable e incluye a Π_1, entonces existe una oración tal que ni ella ni su negación pertenecen a Δ.*

Prueba. Para cada $a \in \mathbb{N}$, sea $ndiag\,(a)$ el número obtenido al concatenar el código del símbolo de negación con el número $diag\,(a)$, ambos en su representación decimal. Es decir: $ndiag\,(a)$ es igual al resultado de concatenar "1000000" con el número $diag\,(a)$.

Consideremos entonces aquella propiedad Q tal que para todo $a \in \mathbb{N}$,

$$Q(a) \qquad \text{si y sólo si} \qquad Or_\Delta\,(\,ndiag\,(a)\,)$$

Además, sea Q' a la sazón aquella propiedad tal que para todo $a \in \mathbb{N}$,

$$Q'\,(a) \qquad \text{si y sólo si} \qquad Or_\Delta\,(\,diag\,(a)\,)$$

Puesto que Δ es, por hipótesis, una teoría recursivamente axiomatizable, la propiedad Or_Δ ha de ser r.e., con lo que aplicando la *Práctica* § 5.23(c) (p. 104), se sigue de inmediato que Q' también es r.e.

Por otra parte, la función $ndiag$ también es obviamente una función recursiva. De lo cual se sigue a su vez igualmente, y por las mismas razones, que también Q es una propiedad r.e.

Por lo tanto, por nuestro lema sobre la representación en Π_1 (§ 6.13, p. 114), tomando el caso $n = 1$, existirá una fórmula $\underline{\alpha} \in \Upsilon_1$ tal que para cualquier $a \in \mathbb{N}$,

$$\text{si} \qquad Q(a) \qquad \text{y} \qquad \sim Q'\,(a) \qquad \text{entonces} \qquad \underline{\alpha\,[s_a]} \in \Pi_1 \tag{1}$$

$$\text{si} \qquad \sim Q(a) \qquad \text{y} \qquad Q'\,(a) \qquad \text{entonces} \qquad \underline{\neg\alpha\,[s_a]} \in \Pi_1 \tag{2}$$

A continuación razonamos como sigue. Por definición de Q' tenemos

$$Q'\,(\sharp\,\underline{\alpha}) \qquad \text{si y sólo si} \qquad Or_\Delta\,(\,diag\,(\sharp\,\underline{\alpha})\,)$$

es decir,

$$Q'\,(\sharp\,\underline{\alpha}) \qquad \text{si y sólo si} \qquad \underline{\alpha\,[s_{\sharp\,\alpha}]} \in \Delta \tag{3}$$

Por su parte, por definición de Q,

$$Q(\sharp\,\underline{\alpha}) \qquad \text{si y sólo si} \qquad Or_\Delta\,(\,ndiag\,(\sharp\,\underline{\alpha})\,)$$

Ahora bien, obviamente el número $ndiag\,(\sharp\,\underline{\alpha})$, es decir, el resultado de anteponer un "1000000" al código de $\underline{\alpha\,[s_{\sharp\alpha}]}$, es precisamente el código de la oración $\underline{\neg\alpha\,[s_{\sharp\,\alpha}]}$. Por lo tanto tenemos:

$$Q(\sharp\,\underline{\alpha}) \qquad \text{si y sólo si} \qquad \underline{\neg\alpha\,[s_{\sharp\alpha}]} \in \Delta \tag{4}$$

Supongamos ahora que $\underline{\alpha\,[s_{\sharp\alpha}]}$ perteneciese a Δ. Entonces sucedería $Q'\,(\sharp\,\underline{\alpha})$, por (3). Como Δ es consistente, sucedería también que $\underline{\neg\alpha\,[s_{\sharp\,\alpha}]} \notin \Delta$, y por tanto $\sim Q(\sharp\,\underline{\alpha})$, por (4). Por consiguiente, aplicando (2) se seguiría que $\underline{\neg\alpha\,[s_{\sharp\alpha}]} \in \Pi_1 \subseteq \Delta$, lo cual es imposible ya que Δ es consistente.

Por tanto la oración $\underline{\alpha\,[s_{\sharp\alpha}]}$ no pertenece a Δ.

Y por otro lado, supongamos ahora que fuese su negación, $\underline{\neg\alpha\,[s_{\sharp\alpha}]}$, la que perteneciese a Δ. Entonces tendríamos $Q(\sharp\,\underline{\alpha})$. De nuevo por la consistencia de Δ se seguiría $\underline{\alpha\,[s_{\sharp\alpha}]} \notin \Delta$, y por tanto $\sim Q'\,(\sharp\,\underline{\alpha})$. Por consiguiente, aplicando (1) se seguiría que $\underline{\alpha\,[s_{\sharp\alpha}]} \in \Pi_1 \subseteq \Delta$, lo cual es imposible por la misma razón que en el caso anterior.

Así que en definitiva, hemos demostrado que la oración $\underline{\neg\alpha\,[s_{\sharp\alpha}]}$ tampoco pertenece a Δ.

§ 6.36. Observación (La versión sintáctica del primer teorema de incompletitud). También la oración construida en esta ocasión, $\underline{\alpha\,[s_{\sharp\alpha}]}$, existe, y concretamente es una oración *falsa* en $\mathscr{N}$, aunque ahora resulta algo más complicado desentrañar qué es lo que tal oración viene indirectamente a decir, a través del método de codificación. Ello depende en gran parte de la prueba del *Lema* § 6.13 que hemos omitido, por lo que resultaría demasiado complicado reconstruir toda la explicación aquí.

§ 6.37. Práctica⋆ (Teorías inconsistentes). Mostrar que una teoría de $\mathscr{A}$ es inconsistente si y sólo si contiene la oración $\underline{0 \neq 0}$.

§ 6.38. Observación (La oración *con*). Puesto que $\underline{0 \neq 0}$ es la oración $\underline{\neg\, 0 = 0}$, es claro que el código de $\underline{0 \neq 0}$ es el resultado de concatenar el código del negador con el código de la oración $\underline{0 = 0}$.

Por otro lado, resulta obvio que el resultado de sustituir en $\underline{0 = 0}$ la variable $\underline{v_1}$ por cualquier otro término, deja a esta oración inalterada, ya que tal variable no ocurre en $\underline{0 = 0}$, ni ligada ni libre. Por tanto, en particular, tampoco alterará a esta oración la substitución $\underline{[v_1/s_{\sharp\,0=0}]}$, lo cual quiere decir que

$$diag\,(\sharp\,\underline{0 = 0}) \qquad = \qquad \sharp\,\underline{0 = 0}$$

Ahora sea Δ cualquier teoría de $\mathscr{A}$ recursivamente axiomatizable. Según la Práctica precedente, Δ será inconsistente si y sólo si contiene la oración $\underline{0 \neq 0}$, es decir, si y sólo si $Or_\Delta(\sharp\,\underline{0 \neq 0})$. De acuerdo con lo que acabamos de ver, esto es equivalente a que se dé

$$Or_\Delta\,(\,ndiag\,(\sharp\,\underline{0 = 0})\,)$$

Y por tanto a que se dé $Q(\sharp\,\underline{0 = 0})$, donde Q es la propiedad definida en la prueba del teorema anterior (§ 6.35).

Por lo tanto tenemos:

$$\Delta \text{ es inconsistente} \qquad \text{si y sólo si} \qquad Q(\sharp\,\underline{0 = 0})$$

O lo que es lo mismo:

$$\Delta \text{ es consistente} \qquad \text{si y sólo si} \qquad \sim Q(\sharp\,\underline{0 = 0})$$

Según se vio en dicha prueba, Q es r.e., con lo que habrá una fórmula $\underline{\beta} \in \Upsilon_1$ que represente a Q en Ω. Así que tendremos

$$\sim Q(\sharp\,\underline{0 = 0}) \qquad \text{si y sólo si} \qquad \underline{\neg\beta\,[s_{\sharp\,0=0}]} \in \Omega$$

y por consiguiente:

$$\Delta \text{ es consistente} \qquad \text{si y sólo si} \qquad \underline{\neg\beta\,[s_{\sharp\,0=0}]} \in \Omega$$

Es decir: Δ es consistente si y sólo si $\underline{\neg\beta\,[s_{\sharp\,0=0}]}$ es una oración verdadera en $\mathscr{N}$.

De esta manera, vemos que $\underline{\neg\beta\,[s_{\sharp\,0=0}]}$ viene a decir, indirectamente, a través del método de codificación y por referencia a la teoría Ω, que Δ es una teoría consistente.

Pues bien, todo ello justifica la siguiente definición.

§ 6.39. Definición (La oración $\underline{con}$). Si Δ es cualquier teoría de $\mathscr{A}$ recursivamente axiomatizable, entonces $\underline{con_\Delta}$ es la correspondiente oración $\underline{\neg\beta\,[s_{\sharp\,0=0}]}$ construida según se ha indicado en la observación precedente.

§ 6.40. Teorema (Segundo teorema de incompletitud). *Sea Δ cualquier teoría recursivamente axiomatizable que incluya a Π, la aritmética de Peano de primer orden. Entonces, si Δ es consistente, la oración $\underline{con_\Delta}$ no pertenece a Δ.*

Prueba. Sea Δ cualquier teoría recursivamente axiomatizable que incluya a la teoría Π. Entonces incluirá naturalmente a Π_1, y se le aplicará la versión sintáctica del primer teorema de incompletitud (§ 6.35). Por tanto, tomando la oración $\underline{\alpha\,[s_{\sharp\,\alpha}]}$ descrita en la prueba de dicho teorema, tenemos en particular que

$$\text{si } \Delta \text{ es consistente, entonces } \underline{\neg\alpha\,[s_{\sharp\alpha}]} \notin \Delta \tag{1}$$

De acuerdo con la definición precedente, resulta inmediato expresar de una manera formal en el lenguaje $\mathscr{A}$ y por referencia a $\mathscr{N}$, el antecedente de este enunciado condicional (1): basta con utilizar la oración $\underline{con_\Delta}$ allí descrita.

En cuanto al consecuente de (1), tampoco es difícil de expresar en $\mathscr{A}$, como vamos a ver. En efecto, el hecho de que $\underline{\neg\alpha\,[s_{\sharp\alpha}]} \notin \Delta$ es obviamente equivalente a que *no se dé* $Or_\Delta\,(\,ndiag\,(\sharp\,\underline{\alpha}))$. Es decir, a que se dé

$$\sim Q(\sharp\,\underline{\alpha})$$

donde Q es también la propiedad referida en la prueba del mencionado *Teorema* § 6.35.

Ahora bien, tal y como comentamos en § 6.38, al ser Q r.e. estará representada en Ω por una fórmula $\underline{\beta} \in \Upsilon_1$. Precisamente la misma fórmula *beta* utilizada para definir la oración $\underline{con_\Delta}$. Por consiguiente, el hecho de que $\underline{\neg\alpha\,[s_{\sharp\alpha}]} \notin \Delta$ será equivalente a que la oración $\underline{\neg\beta\,[s_{\sharp\alpha}]}$ pertenezca a Ω, es decir: a que $\underline{\neg\beta\,[s_{\sharp\alpha}]}$ sea una oración verdadera en $\mathscr{N}$.

De todo ello se sigue que la oración

$$\underline{con_\Delta \;\rightarrow\; \neg\beta\,[s_{\sharp\alpha}]} \tag{2}$$

esto es, la oración

$$\underline{\neg\beta\,[s_{\sharp\,0=0}] \;\rightarrow\; \neg\beta\,[s_{\sharp\alpha}]}$$

expresa, por referencia a $\mathscr{N}$, justamente el hecho descrito por (1).

Es más, puesto que (1) es efectivamente el caso, como se probó en § 6.35, la oración (2) será verdadera en $\mathscr{N}$, y por lo tanto pertenecerá a Ω.

Y en realidad, (2) no sólo pertenece a Ω, sino también a la propia Π, la aritmética de Peano de primer orden. La comprobación de este hecho requiere simplemente examinar la larga cadena de razonamientos que nos llevó a poder establecer el *Teorema* § 6.35, reproduciéndolos en $\mathscr{A}$ a través del método de codificación, y deduciéndolos uno a uno de los axiomas de la teoría formal Π.

Dicha tarea no es demasiado instructiva, pero sí larguísima y extraordinariamente engorrosa, por lo que la omitimos sin dudarlo.

Así que en definitiva, tenemos que

$$\underline{con_\Delta \ \rightarrow \ \neg\beta\,[s_{\sharp\alpha}]} \in \Pi$$

A continuación, aunque nosotros hemos omitido asimismo la prueba del lema sobre la representación en Π_1 (*Lema* § 6.13, p. 114), se establece en dicha prueba que, para estas mismas fórmulas $\underline{\alpha}$ y $\underline{\beta}$ que estamos manejando aquí, y dado cualquier $a \in \mathbb{N}$, tenemos

$$\underline{\neg\beta\,[s_a]} \vdash \underline{\neg\alpha\,[s_a]}$$

Con lo que en particular,

$$\underline{\neg\beta\,[s_{\sharp\,\alpha}]} \vdash \underline{\neg\alpha\,[s_{\sharp\alpha}]}$$

Y al ser Π una teoría, tendremos

$$\underline{con_\Delta \ \rightarrow \ \neg\alpha\,[s_{\sharp\alpha}]} \in \Pi \tag{3}$$

Y por último, como Δ es por hipótesis una teoría consistente, entonces, una vez más por el *Teorema* § 6.35 tenemos que $\underline{\neg\alpha\,[s_{\sharp\alpha}]} \notin \Delta$, y dado que $\Pi \subseteq \Delta$, de (3) se sigue inmediatamente que $\underline{con_\Delta} \notin \Delta$.

§ 6.41. Observación (El segundo teorema de incompletitud). El segundo teorema de incompletitud se aplica a cualquier teoría formal suficientemente fuerte, tanto del lenguaje $\mathscr{A}$ como de otros lenguajes formales. Todo lo que se requiere es que la teoría en cuestión sea recursivamente axiomatizable, y que incluya a Π, o a alguna contrapartida de Π en el lenguaje de esa teoría.

En su momento, nosotros vimos que Π era una teoría correcta con respecto al modelo estándar $\mathscr{N}$ (§ 6.19, p. 115), de lo que se deriva de inmediato que es también una teoría consistente.

Ahora comprobamos, sin embargo, que la consistencia de Π no puede ser establecida por medio de argumentos que, una vez formalizados, sean deducibles de sus propios axiomas. Es decir: que Π no es lo bastante fuerte como para implicar su propia consistencia.

Y lo mismo sucederá con cualquier otra teoría formal recursivamente axiomatizable que incluya a Π. Es decir: podremos dar una prueba de consistencia de Π por medio de una formalización de la teoría de conjuntos, por ejemplo, utilizando la representación de los números naturales como conjuntos a la que aludimos en § 1.35 (p. 34). Pero entonces nos resultará imposible dar una prueba de consistencia de la nueva teoría utilizada, *dentro* de ella misma, esto es, dentro de esa nueva teoría. Habría que acudir a otra teoría aún más fuerte, a la cual le sucedería lo mismo, etc. Y el proceso no terminaría nunca.

Ello hace, en definitiva, que la tarea de justificar la consistencia de Π sin hacer referencia a ninguna interpretación suya sea virtualmente imposible. Lo cual a su vez constituye un obstáculo muy serio para la concepción formalista, que lo que intenta es precisamente eliminar dicha referencia. Sin apoyarnos en la existencia de la estructura abstracta $\mathscr{N}$, no hay medios efectivos de establecer que Π es una teoría consistente, con lo cual dicha teoría pierde prácticamente todo su interés.

Por último, por lo que respecta otra vez a nuestra comparación con los ordenadores, resulta patente que, en la medida en que su programa de funcionamiento tiene que asimilarse de algún modo a una teoría formal, estará siempre constreñido por la limitación señalada a la hora de establecer formalmente su propia consistencia. Ahora bien: tampoco nosotros podemos establecer *formalmente* la consistencia de Π, o de teorías como Π: si creemos en esa consistencia es por otras razones, como que creemos tener un conocimiento de la estructura abstracta que representa, es decir, el modelo $\mathscr{N}$, creemos tener la capacidad de referirnos a ella, etc.

Es por ello que, en mi opinión, el resultado precedente, que sirve de colofón a esta asignatura, sólo puede apuntar a una diferencia fundamental entre los ordenadores y la estructura de la mente humana, si mantenemos de entrada la posición de negar a los primeros toda capacidad de referencia intencional, o de conocimiento con contenido, mediante los cuales también ellos pudieran conocer y referirse a la estructura $\mathscr{N}$. Y en cualquier caso, aun manteniendo de entrada esa diferencia fundamental entre ellos y nosotros, el segundo teorema de incompletitud no aportaría, bajo mi punto de vista, ningún argumento esencial en favor de la misma.

Todo ello no obstante, hay que advertir que este es un asunto ciertamente polémico, sobre el cual hay una clara falta de consenso entre los distintos autores. Para un ilustrativo cruce de opiniones al respecto, pueden consultarse, por ejemplo, Nagel y Newman, *El teorema de Gödel*, en esp. pp. 117–121, y Hilary Putnam, en la compilación Turing, Putnam, Davidson, *Mentes y máquinas*, en esp. p. 70.

Bibliografía general

Agazzi, E., *La lógica simbólica* (4ª ed. cast.), Barcelona: Herder, 1986.

Alchourrón, C. E. (ed.), *Lógica*, Madrid: Trotta, 1995.

Badesa, C., I. Jané y R. Jansana, *Elementos de lógica formal*, Barcelona: Ariel, 1998.

Barwise, K. J. (ed.), *Handbook of Mathematical Logic*, Amsterdam: North-Holland, 1977 (reimp. 1982).

Bell, J. L., y M. Machover, *A Course in Mathematical Logic*, Amsterdam: North-Holland, 1977 (reimp. 1997).

Boolos, G. S., J. P. Burgess y R. C. Jeffrey, *Computability and Logic* (4ª ed.), Cambridge: Cambridge University Press, 2003.

Braun, R., *Elementos de lógica simbólica*, Lima: Universidad de Lima, 1995.

Chaiting, G. J., *The Limits of Mathematics: A Course on Information Theory and the Limits of Formal Reasoning*, Singapur: Springer, 1998.

Chang, C. C., y H. J. Keisler, *Model Theory* (3ª ed.), Amsterdam: North-Holland, 1990 (reimp. 1998).

Cohen, D. E., *Computability and Logic*, Chichester (West Sussex): Ellis Horward, 1987.

Crossley, J. N., C. J. Ash, C. J. Brickhill, J. C. Stilwell y N. H. Williams, *¿Qué es la lógica matemática?*, Madrid: Tecnos, 1986.

Davis, M., *La computadora universal: de Leibniz a Turing*, Madrid: Debate, 2002.

Díaz Estévez, E., *El teorema de Gödel*, Pamplona: Eunsa, 1975.

Drake, F. R., y D. Singh, *Intermediate Set Theory*, Chichester (West Sussex): Wiley & Sons, 1996.

Ebbinghaus, H. D., J. Flum y W. Thomas, *Mathematical Logic* (2ª ed.), Nueva York: Springer, 1994.

Enderton, H. B., *Una introducción matemática a la lógica*, México D.F.: Universidad Nacional Autónoma de México, 1987.

Ershov, Yu., y E. Paliutin, *Lógica matemática*, Moscú: Mir, 1990.

Ferrater Mora, J., *Diccionario de Filosofía* (4 vols., ed. actualizada), Barcelona: Ariel, 1994.

Fraenkel, A. A., Y. Bar-Hillel y A. Levy, *Foundations of Set Theory* (2ª ed.), Amsterdam: North-Holland, 1973 (reimp. 1984).

Gallardo López, D., P. Arqués Corrales e I. Lesta Pelayo, *Introducción a la teoría de la computabilidad*, Alicante: Universidad de Alicante, 1998.

Gensler, H. J., *Gödel's Theorem Simplified*, Lauham (MD.): University Press of America, 1984.

Gödel, K., *Obras completas* (2ª ed.), Madrid: Alianza, 1989.

Hajek, P., y P. Pudlak, *Metamathematics of First-Order Arithmetic*, Berlín: Springer, 1993.

Halmos, P. R., *Teoría intuitiva de conjuntos*, México D.F.: Continental, 1965 (reimp. 1984).

Hamilton, A. G., *Lógica para matemáticos*, Madrid: Paraninfo, 1981.

Heijenoort, J. van (ed.), *From Frege to Gödel: A Sourcebook in Mathematical Logic, 1879–1931*, Cambridge (MA.): Harvard University Press, 1967 (reimp. 1999).

Hermes, H., *Introducción a la teoría de la computabilidad*, Madrid: Tecnos, 1984.

Hodges, W., *Model Theory*, Cambridge: Cambridge University Press, 1993.

Hodges, W., *A Shorter Model Theory*, Cambridge: Cambridge University Press, 1997.

Hunter, G., *Metalógica*, Madrid: Paraninfo, 1981.

Ladrière, J., *Limitaciones internas de los formalismos*, Madrid: Tecnos, 1969.

Lipschutz, S., *Teoría de conjuntos y temas afines*, en México D.F.: McGraw-Hill, 1969.

Liz Gutiérrez, M., *Teoría intuitiva de conjuntos y lógica clásica de proposiciones: Ejercicios*, La Laguna: Universidad de La Laguna, 1990.

Lorenzo, J. de, *Iniciación a la teoría intuitiva de conjuntos*, Madrid: Tecnos, 1972.

Machover, M., *Set Theory, Logic and their Limitations*, Cambridge: Cambridge University Press, 1996.

Manin, Yu. I., *A Course in Mathematical Logic*, Berlín: Springer, 1977 (reimp. 1984).

Manzano, M., *Teoría de modelos*, Madrid: Alianza, 1989 (trad. al inglés en Oxford: Clarendon Press, 1999).

Marraud González, H., *Teoría de modelos elemental*, Cantoblanco (Madrid): Universidad Autónoma de Madrid, 1990.

Mendelson, E., *Introduction to Mathematical Logic* (4ª ed.), Londres: Chapman & Hall, 1997.

Mosterín, J., *Teoría axiomática de conjuntos* (2ª ed.), Barcelona: Ariel, 1980.

Mosterín, J., *Lógica de primer orden* (3ª ed.), Barcelona: Ariel, 1983.

Mosterín, J., *Los lógicos*, Madrid: Espasa Calpe, 2000.

Nagel, E., y J. R. Newman, *El teorema de Gödel*, Madrid: Tecnos, 1970 (reimp. 1994).

Odifreddi, P. G., *Classical Recursion Theory* (2 vols.), Amsterdam: Elsevier, 1989–1999.

Peña, L., *Rudimentos de Lógica Matemática*, Madrid: Consejo Superior de Investigaciones Científicas, 1991.

Rogers, H. Jr., *Theory of Recursive Functions and Effective Computability* (2ª ed.), Nueva York: McGraw-Hill, 1987.

Roitman, J., *Introduction to Modern Set Theory*, Nueva York: Wiley, 1990.

Sacks, G. E., *General Recursion Theory*, Berlín: Springer, 1990.

Sacristán Luzón, M., *Introducción a la lógica y al análisis formal* (2ª ed.), Barcelona: Ariel, 1996.

Sancho San Román, J., *Lógica matemática y computabilidad*, Madrid: Ediciones Díaz de Santos, 1990.

Sanmartín Esplugues, J., *Una introducción constructiva a la teoría de modelos* (2ª ed.), Madrid: Tecnos, 1983.

Shanker, S. G. (ed.), *Gödel's Theorem in Focus*, Londres: Routledge, 1990.

Shoenfield, J. R., *Mathematical Logic*, Londres: Addison-Wesley, 1967 (reimp. en Natick (MA.): A. K. Peters, 2000).

Shoenfield, J. R., *Degrees of Unsolvability*, Amsterdam: North-Holland, 1971.

Shoenfield, J. R., *Recursion Theory*, Berlín: Springer, 1993 (reimp. en Natick (MA.): A. K. Peters, 2000).

Smullyan, R. M., *Lógica de primer orden*, Madrid: Tecnos, 1983.

Smullyan, R. M., *Gödel's Incompleteness Theorems*, Oxford: Clarendon Press, 1992.

Takeuti, G., *Proof Theory* (2ª ed.), Amsterdam: North-Holland, 1987.

Troelstra, A. S., y H. Schwichtenberg, *Basic Proof Theory* (2ª ed.), Cambridge: Cambridge University Press, 2000.

Turing, A. M, H. Putnam y D. Davidson, *Mentes y máquinas*, Madrid: Tecnos, 1985.

Zalabardo, J. L., *Introducción a la teoría de la lógica*, Madrid: Alianza, 2002 (pub. orig. en inglés, en Oxford: Westview Press, 2000).

Índice general

La inducción matemática

Módulo 2: *Elementos de lógica de primer orden*

Términos

Fórmulas

Modelos

El teorema de Tarski

Teorías recursivamente axiomatizables

Módulo 6: *El teorema de Church y los teoremas de Gödel*

La aritmética de Peano de primer orden