Lógica y Teoría Descriptiva de Conjuntos

Los Del DGIIM, losdeldgiim.github.io

Doble Grado en Ingeniería Informática y Matemáticas
Universidad de Granada

Lógica y Teoría Descriptiva de Conjuntos

Los Del DGIIM, losdeldgiim.github.io

José Juan Urrutia Milán

Granada, 2025

Índice general

El presente documento es un resumen del microcredencial de "Lógica y Teoría Descriptiva de Conjuntos", que recoge los principales conceptos que se impartieron en el mismo. Si cursa el microcredencial se recomienda ver los recursos proporcionados por el profesorado. Si está cursando actualmente la asignatura de "Lógica y Métodos Discretos" del grado de Informática, los dos primeros capítulos pueden serle de gran ayuda.

A lo largo del curso trabajaremos en $\mathbb{Z}_2$, por lo que se recomienda al lector repasar los apuntes de Álgebra I en caso de no estar familiarizado con dicho cuerpo.

1. Lógica Proposicional

Consideraremos un conjunto finito de proposiciones atómicas, que serán para nosotros enunciados indivisibles. Nos interesará la veracidad o falsedad de cada una de estas proposiciones. Consideraremos sobre estas las conectivas $\neg$, $\wedge$, $\vee$, $\rightarrow$ y $\leftrightarrow$. De esta forma, somos capaces de definir lo que es una proposición en nuestro lenguaje.

Definición 1.1 (Proposición)**.** Definimos las proposiciones de forma recursiva[1]:

1. Las proposiciones atómicas son proposiciones.
2. Si α y β son proposiciones, también lo son:

$$\neg\alpha,\ \alpha \wedge \beta,\ \alpha \vee \beta,\ \alpha \rightarrow \beta,\ \alpha \leftrightarrow \beta$$

3. No hay más proposiciones que las que se puedan obtener siguiendo una secuencia finita de pasos a partir de las enunciadas.

1.1. Semántica

Una vez definida lo que es una proposición, pasamos a lo que nos interesa, asignar un valor de verdad o de falsedad a cada una de las proposiciones que nos encontremos. Para ello, consideraremos una aplicación del conjunto de las proposiciones en $\mathbb{Z}_2$, e interpretaremos el valor de 0 como falso y el valor de 1 como verdad.

Definición 1.2 (Interpretación)**.** Sea $\mathcal{P}$ el conjunto de todas las proposiciones de un lenguaje proposicional, una interpretación sobre el mismo es una aplicación $I : \mathcal{P} \rightarrow \mathbb{Z}_2$ que verifica:

1. $I(\neg a) = 1 + I(a)$.
2. $I(a \wedge b) = I(a)I(b)$.
3. $I(a \vee b) = I(a) + I(b) + I(a)I(b)$.
4. $I(a \rightarrow b) = 1 + I(a) + I(a)I(b)$.
5. $I(a \leftrightarrow b) = 1 + I(a) + I(b)$.

Para cualesquiera proposiciones $a, b \in \mathcal{P}$.

[1] Algo que será habitual en este curso.

Observación. Observemos que, gracias a la naturaleza recursiva de las interpretaciones, basta dar un valor de $\mathbb{Z}_2$ a cada proposición atómica para obtener una interpretación: conocidos los valores de las proposiciones atómicas conocemos el valor de cualquier proposición y viceversa.

Definición 1.3. Sea α y β dos proposiciones de forma que $I(\alpha) = I(\beta)$ para cualquier interpretación I, entonces escribiremos que $\alpha \equiv \beta$ y podemos decir que α y β son semánticamente equivalentes.

Definición 1.4. Sea α una proposición:

- Si existe una interpretación I de forma que $I(\alpha) = 1$, diremos que p es **satisfacible**.
- Si existe una interpretación I de forma que $I(\alpha) = 0$, diremos que p es **refutable**.
- Si $I(\alpha) = 1$ para cualquier interpretación I, diremos que p es una **tautología**.
- Si $I(\alpha) = 0$ para cualquier interpretación I, diremos que p es una **contradicción**.

Definición 1.5 (Consecuencia lógica)**.** Sea $\Gamma \cup \{p\}$ un conjunto de proposiciones, decimos que p es consecuencia lógica de Γ (notado por $\Gamma \vDash p$), si dada una interpretación I, siempre que se tenga que $I(\gamma) = 1$ para cualquier $\gamma \in \Gamma$, entonces se tiene que $I(p) = 1$.

Notación. Por comodidad, si p es una proposición de forma que $\emptyset \vDash p$, entonces notaremos:

$$\vDash p$$

Notemos que en este caso p es una tautología, ya que estamos diciendo que $I(p) = 1$ para cualquier[2] interpretación I.

Proposición 1.1. *Se verifica que* $\Gamma \vDash p$ *si y solo si* $(1 + I(p)) \prod_{\gamma \in \Gamma} I(\gamma) = 0$.

Demostración. Veamos las dos implicaciones:

$\Longrightarrow$) Sea I una interpretación:

- Si existe un $\gamma \in \Gamma$ de forma que $I(\gamma) = 0$, entonces tenemos el resultado.
- En caso contrario, tendremos que $I(\gamma) = 1$ para cualquier $\gamma \in \Gamma$. En dicho caso, como $\Gamma \vDash p$, se tendrá que $I(p) = 1$, por lo que:

$$1 + I(p) = 0 \Longrightarrow (1 + I(p)) \prod_{\gamma \in \Gamma} I(\gamma) = 0$$

$\Longleftarrow$) Sea I una interpretación que verifica $I(\gamma) = 1$ para cualquier $\gamma \in \Gamma$, como $\mathbb{Z}_2$ es un dominio de integridad, de $(1 + I(p)) \prod_{\gamma \in \Gamma} I(\gamma) = 0$ deducimos que $I(p) + 1 = 0$, por lo que $I(p) = 1$ y entonces se tiene que $\Gamma \vDash p$.

[2] Cualquiera que haga ciertos todos los elementos del vacío.

□

Teorema 1.2 (de la deducción)**.** *Sea* $\Gamma \cup \{\alpha, \beta\}$ *un conjunto de proposiciones, equivalen:*

1. $\Gamma \vDash \alpha \to \beta$
2. $\Gamma \cup \{\alpha\} \vDash \beta$

Demostración. Demostramos las dos implicaciones:

1) $\Longrightarrow$ 2) Sea I una interpretación de forma que $I(\alpha) = 1$ y que $I(\gamma) = 1$ para todo $\gamma \in \Gamma$, entonces (por 1) deducimos que $I(\alpha \to \beta) = 1$, luego:

$$1 = I(\alpha \to \beta) = 1 + \cancelto{1}{I(\alpha)} + \cancelto{1}{I(\alpha)} I(\beta) = 1 + 1 + I(\beta) = I(\beta)$$

2) $\Longrightarrow$ 1) Sea I una interpretación de forma que $I(\gamma) = 1$ para todo $\gamma \in \Gamma$:

- Si $I(\alpha) = 0$, entonces:
$$I(\alpha \to \beta) = 1 + I(\alpha) + I(\alpha)I(\beta) = 1$$
Por lo que se tiene 1.
- Si $I(\alpha) = 1$, como $\Gamma \cup \{\alpha\} \vDash \beta$, entonces $I(\beta) = 1$, por lo que:
$$I(\alpha \to \beta) = 1 + I(\alpha) + I(\alpha)I(\beta) = 1 + 1 + 1 = 1$$

□

Ejemplo. Demostraremos ahora que varias proposiciones son tautologías:

$\vDash \alpha \to \alpha$

Por el Teorema de la deducción (**??**), $\vDash \alpha \to \alpha$ es equivalente a ver que $\{\alpha\} \vDash \alpha$. En efecto, sea I una interpretación de forma que $I(\alpha) = 1$, tenemos que $I(\alpha) = 1$.

$\vDash \alpha \to (\beta \to \alpha)$

Por el Teorema de la deducción, es equivalente ver que $\{\alpha\} \vDash \beta \to \alpha$; que nuevamente por el Teorema de la deducción es equivalente ver que $\{\alpha, \beta\} \vDash \alpha$. En efecto, sea I una interpretación de forma que $I(\alpha) = I(\beta) = 1$, entonces $I(\alpha) = 1$.

$\vDash (\alpha \to (\beta \to \gamma)) \to ((\alpha \to \beta) \to (\alpha \to \gamma))$

Por el Teorema de la deducción aplicado 3 veces, es equivalente ver que:

$$\{\alpha \to (\beta \to \gamma), \alpha \to \beta, \alpha\} \vDash \gamma$$

Sea I una interpretación de forma que:

$$1 = I(\alpha \to (\beta \to \gamma)) = I(\alpha \to \beta) = I(\alpha)$$

Entonces:

$$\begin{aligned} &1 = I(\alpha \to \beta) = 1 + I(\alpha) + I(\alpha)I(\beta) = 1 + 1 + I(\beta) = I(\beta) \Longrightarrow I(\beta) = 1 \\ &1 = I(\alpha \to (\beta \to \gamma)) = 1 + I(\alpha) + I(\alpha)I(\beta \to \gamma) \\ &\quad = 1 + I(\alpha) + I(\alpha)(1 + I(\beta) + I(\beta)I(\gamma)) = 1 + 1 + 1(1 + 1 + I(\gamma)) \\ &\quad = I(\gamma) \Longrightarrow \underline{I(\gamma) = 1} \end{aligned}$$

$\vDash (\neg\alpha \to \neg\beta) \to ((\neg\alpha \to \beta) \to \alpha)$

Por el Teorema de la deducción aplicado 2 veces, es equivalente ver que:

$$\{\neg\alpha \to \neg\beta, \neg\alpha \to \beta\} \vDash \alpha$$

Sea I una interpretación de forma que:

$$1 = I(\neg\alpha \to \neg\beta) = 1 + I(\neg\alpha) + I(\neg\alpha)I(\neg\beta)$$
$$1 = I(\neg\alpha \to \beta) = 1 + I(\neg\alpha) + I(\neg\alpha)I(\beta)$$

Entonces (sumando):

$$0 = I(\neg\alpha \to \neg\beta) + I(\neg\alpha \to \beta) = I(\neg\alpha)(I(\neg\beta) + I(\beta)) \overset{(*)}{=} I(\neg\alpha)$$

Donde en $(*)$ hemos usado que $I(\neg\beta) = 1 + I(\beta) \Longrightarrow I(\neg\beta) + I(\beta) = 1$.

Como $I(\neg\alpha) = 0$, se tiene que $I(\alpha) = 1$, como queríamos demostrar.

Definición 1.6. Sea Γ un conjunto de proposiciones, decimos que Γ es **inconsistente** si para toda interpretación I existe $\gamma \in \Gamma$ de forma que $I(\gamma) = 0$.

Proposición 1.3. *Sea $\Gamma \cup \{\alpha\}$ un conjunto de proposiciones, equivalen:*

1. $\Gamma \vDash \alpha$.
2. $\Gamma \cup \{\neg\alpha\}$ *es inconsistente.*

Demostración. Demostramos las dos implicaciones:

1) $\Longrightarrow$ 2) Sea I una interpretación:

- Si existe un $\gamma \in \Gamma$ de forma que $I(\gamma) = 0$, entonces Γ es inconsistente, de donde $\Gamma \cup \{\neg\alpha\}$ también lo es.
- Si $I(\gamma) = 1$ para cualquier $\gamma \in \Gamma$, aplicando que $\Gamma \vDash \alpha$ deducimos que $I(\alpha) = 1 \Longrightarrow I(\neg\alpha) = 1 + I(\alpha) = 0$, por lo que $\Gamma \cup \{\neg\alpha\}$ es inconsistente.

2) $\Longrightarrow$ 1) Sea I una interpretación de forma que $I(\gamma) = 1$ para cualquier $\gamma \in \Gamma$, como $\Gamma \cup \{\neg\alpha\}$ es inconsistente, deducimos que $I(\neg\alpha) = 0$, luego $I(\alpha) = 1$. □

1.2. Demostraciones

Definición 1.7 (Demostración)**.** Sean $\mathcal{A}$ y $\Gamma \cup \{p\}$ dos conjuntos de proposiciones (nos referiremos al conjunto $\mathcal{A}$ como "conjunto de axiomas" y a Γ como "conjunto de hipótesis"), una demostración de p a partir de Γ (notado por $\Gamma \vdash p$) es una secuencia de proposiciones $\alpha_1, \alpha_2, \ldots, \alpha_n$ de forma que $\alpha_n = p$ y se verifica para todo i menor o igual que n:

- bien $\alpha_i \in \mathcal{A} \cup \Gamma$.
- bien existen j, k naturales con $j < k < i$ siendo $\alpha_k = \alpha_j \to \alpha_i$.

  En este caso, diremos que se tiene α_i por modus ponens de j y k.

Notación. Si p es una proposición de forma que $\emptyset \vdash p$, podremos notar $\vdash p$ y diremos que p es un teorema.

Ejemplo. Como ejemplo de demostración, veamos que $\{\alpha, \alpha \to \beta\} \vdash \beta$ (regla conocida como "Modus ponens"). Para ello, consideramos:

$$\alpha_1 = \alpha$$
$$\alpha_2 = \alpha \to \beta$$
$$\alpha_3 = \beta$$

Como vemos, es una demostración de β a partir de $\{\alpha, \alpha \to \beta\}$ porque $\alpha_1, \alpha_2, \alpha_3$ son proposiciones, $\alpha_3 = \beta$ y:

- $\alpha_1 \in \Gamma$.
- $\alpha_2 \in \Gamma$.
- $1, 2 < 3$ y $\alpha_2 = \alpha_1 \to \alpha_3$.

Notación. Para abreviar las demostraciones, a partir de ahora no daremos una secuencia numerada de proposiciones $\alpha_1, \ldots, \alpha_n$, sino que numeraremos los pasos de la demostración y entenderemos que para formalizarla totalmente debemos coger como α_i el paso $i-$ésimo de la demostración.

Más aún, para no pararnos a comprobar las condiciones abstractas que han de cumplir cada una de las propiedades de la demostrción, incluiremos junto a los pasos de la demostración un comentario sobre por qué dicho paso es válido.

Con esta notación, la demostración de $\{\alpha, \alpha \to \beta\} \vdash \beta$ quedaría de la forma:

1. α es una hipótesis.
2. $\alpha \to \beta$ es una hipótesis.
3. β por Modus Ponens de 1 y 2.

Finalmente, como conjunto $\mathcal{A}$ de axiomas, consideraremos:

$$\mathcal{A} = \mathcal{A}_1 \cup \mathcal{A}_2 \cup \mathcal{A}_3$$

Con:

$$\mathcal{A}_1 = \{\alpha \to (\beta \to \alpha) : \alpha, \beta \text{ son proposiciones}\}$$
$$\mathcal{A}_2 = \{(\alpha \to (\beta \to \gamma)) \to ((\alpha \to \beta) \to (\alpha \to \gamma)) : \alpha, \beta, \gamma \text{ son proposiciones}\}$$
$$\mathcal{A}_3 = \{(\neg\alpha \to \neg\beta) \to ((\neg\alpha \to \beta) \to \alpha) : \alpha, \beta \text{ son proposiciones}\}$$

Ejemplo. Ejemplos de algunas demostraciones:

- $\{\alpha\} \vdash \beta \to \alpha$
  1. $\alpha \to (\beta \to \alpha) \in \mathcal{A}_1$
  2. α es una hipótesis

3. $\beta \to \alpha$ Modus ponens de 1 y 2.

- $\vdash \alpha \to \alpha$

  1. $(\alpha \to ((\alpha \to \alpha) \to \alpha)) \to ((\alpha \to (\alpha \to \alpha)) \to (\alpha \to \alpha)) \in \mathcal{A}_2$
  2. $\alpha \to ((\alpha \to \alpha) \to \alpha) \in \mathcal{A}_1$
  3. $(\alpha \to (\alpha \to \alpha)) \to (\alpha \to \alpha)$ Modus ponens de 1 y 2
  4. $\alpha \to (\alpha \to \alpha) \in \mathcal{A}_1$
  5. $\alpha \to \alpha$ Modus ponens de 3 y 4

Teorema 1.4 (de Herbrand o de la deducción)**.** *Sea* $\Gamma \cup \{\alpha, \beta\}$ *un conjunto de proposiciones, equivalen:*

1. $\Gamma \vdash \alpha \to \beta$

2. $\Gamma \cup \{\alpha\} \vdash \beta$

Demostración. Demostramos las dos implicaciones:

1) $\Longrightarrow$ 2) Como $\Gamma \vdash \alpha \to \beta$, podemos construir una demostración de n pasos de la proposición $\alpha \to \beta$ a partir de Γ. En cuyo caso, podemos añadir 2 pasos más a su demostración, de forma que:

1. ...
   ⋮
n. $\alpha \to \beta$

$n+1$. α es hipótesis

$n+2$. β por Modus ponens de n y $n+1$

Como en los n primeros pasos solo hemos usado como hipótesis Γ, hemos conseguido demostrar en $n+2$ pasos que $\Gamma \cup \{\alpha\} \vdash \beta$.

2) $\Longrightarrow$ 1) Como $\Gamma \cup \{\alpha\} \vdash \beta$, podemos obtener una demostración β a partir de $\Gamma \cup \{\alpha\}$ de n pasos: $\beta_1, \ldots, \beta_n$ (con $\beta_n = \beta$). Por inducción sobre n (el número de pasos de la demostración):

- Si $n = 1$: Como $\Gamma \cup \{\alpha\} \vdash \beta$ gracias a la demostración $\beta_1 = \beta$, distinguimos casos:
  (a) $\beta_1 \in \mathcal{A}$. En dicho caso, podemos considerar la demostración:
    1. $\beta_1 \in \mathcal{A}$
    2. $\beta_1 \to (\alpha \to \beta_1) \in \mathcal{A}_1$
    3. $\alpha \to \beta_1$ por Modus ponens de 1 y 2

  Y con esto tenemos que $\Gamma \vdash \alpha \to \beta$.
  (b) $\beta_1 \in \Gamma$. En dicho caso, podemos considerar una demostración similar al caso anterior:
    1. $\beta_1 \in \Gamma$
    2. $\beta_1 \to (\alpha \to \beta_1) \in \mathcal{A}_1$

3. $\alpha \to \beta_1$ por Modus ponens de 1 y 2

Y con esto también tenemos que $\Gamma \vdash \alpha \to \beta$.

(c) $\beta_1 = \alpha$. En dicho caso, podemos copiar la demostración de $\vdash \beta \to \beta$ del ejemplo anterior, llegando a que $\Gamma \vdash \alpha \to \beta$.

- En el paso de inducción, supuesto que de $\Gamma \cup \{\alpha\} \vdash \beta_m$ podemos deducir que $\Gamma \vdash \alpha \to \beta_m$ para todo $m \leqslant n$, suponemos ahora que $\Gamma \cup \{\alpha\} \vdash \beta_{n+1}$ y queremos ver que $\Gamma \vdash \alpha \to \beta_{n+1}$.

  En dicho caso, supuesto que $\beta_{m+1} \notin \mathcal{A} \cup \Gamma \cup \{\alpha\}$ (ya que si no la demostración es análoga al caso $n = 1$), la única posibilidad es que hayan de existir $i, j < n + 1$ con $\beta_i = \gamma$ y $\beta_j = \gamma \to \beta_{m+1}$.

  Si ahora consideramos los i primeros pasos de la demostración, tenemos que $\Gamma \cup \{\alpha\} \vdash \gamma$ y si consideramos los j primeros pasos, tenemos que $\Gamma \cup \{\alpha\} \vdash \gamma \to \beta_{n+1}$. Por hipótesis de inducción, como $i, j < n + 1$, tenemos que $\Gamma \vdash \alpha \to \gamma$ y que $\Gamma \vdash \alpha \to (\gamma \to \beta_{n+1})$. En este momento, podemos realizar la demostración (con hipótesis Γ):

1. ...
   ⋮
p. $\alpha \to \gamma$
$p+1$. ...
   ⋮
q. $\alpha \to (\gamma \to \beta_{n+1})$
$q+1$. $(\alpha \to (\gamma \to \beta_{n+1})) \to ((\alpha \to \gamma) \to (\alpha \to \beta_{n+1})) \in \mathcal{A}_2$
$q+2$. $(\alpha \to \gamma) \to (\alpha \to \beta_{n+1})$ por Modus ponens de q y $q+1$.
$q+3$. $\alpha \to \beta_{n+1}$ por Modus ponens de p y $q+2$.

□

1.2.1. Resultados útiles a la hora de realizar demostraciones

Proposición 1.5 (Regla de reducción al absurdo clásica)**.** *Sea* $\Gamma \cup \{\alpha, \beta\}$ *un conjunto de proposiciones: si* $\Gamma \cup \{\neg\alpha\} \vdash \beta$ *y* $\Gamma \cup \{\neg\alpha\} \vdash \neg\beta$*, entonces* $\Gamma \vdash \alpha$*.*

Demostración. Supuesto que $\Gamma \cup \{\neg\alpha\} \vdash \beta$ y que $\Gamma \cup \{\neg\alpha\} \vdash \neg\beta$, por el Teorema de Herbrand (**??**), se tiene que $\Gamma \vdash \neg\alpha \to \beta$ y que $\Gamma \vdash \neg\alpha \to \neg\beta$. En dicho caso:

1. ...
   ⋮
p. $\neg\alpha \to \neg\beta$
$p+1$. ...
   ⋮
q. $\neg\alpha \to \beta$
$q+1$. $(\neg\alpha \to \neg\beta) \to ((\neg\alpha \to \beta) \to \alpha) \in \mathcal{A}_3$

$q+2$. $((\neg\alpha \to \beta) \to \alpha)$ por Modus ponens de $q+1$ y p.

$q+3$. α por Modus ponens de $q+2$ y q.

Como desde el paso 1 hasta el q solo hemos usado como hipótesis Γ, deducimos que $\Gamma \vdash \alpha$. □

Proposición 1.6 (Leyes de silogismo o transitividad de la flecha)**.** *Sean α, β y γ proposiciones, se verifican:*

1. $\vdash (\alpha \to \beta) \to ((\beta \to \gamma) \to (\alpha \to \gamma))$
2. $\vdash (\beta \to \gamma) \to ((\alpha \to \beta) \to (\alpha \to \gamma))$

Demostración. Demostraremos la primera y dejamos la segunda como ejercicio. Para ello, aplicando el Teorema de Herbrand 3 veces, llegamos a que 1 es equivalente a ver que:

$$\{\alpha \to \beta, \beta \to \gamma, \alpha\} \vdash \gamma$$

Para ello, nos sirve con la demostración:

1. $\alpha \to \beta$ es una hipótesis
2. α es una hipótesis
3. β por Modus ponens de 1 y 2
4. $\beta \to \gamma$ es una hipótesis
5. γ por Modus ponens de 3 y 4

□

Corolario 1.6.1 (Regla del silogismo)**.** *Sea $\Gamma \cup \{\alpha, \beta\}$ un conjunto de proposiciones, si $\Gamma \vdash \alpha \to \beta$ y $\Gamma \vdash \beta \to \gamma$, entonces $\Gamma \vdash \alpha \to \gamma$.*

Proposición 1.7 (Ley de conmutación de premisas)**.** *Sean α, β y γ proposiciones:*

$$\vdash (\alpha \to (\beta \to \gamma)) \to (\beta \to (\alpha \to \gamma))$$

Demostración. Aplicando el Teorema de Herbrand 3 veces, es equivalente a ver que:

$$\{\alpha \to (\beta \to \gamma), \beta, \alpha\} \vdash \gamma$$

Para ello, nos sirve con:

1. $\alpha \to (\beta \to \gamma)$ es una hipótesis
2. α es una hipótesis
3. $\beta \to \gamma$ por Modus ponens de 1 y 2
4. β es una hipótesis
5. γ por Modus ponens de 3 y 4

□

Corolario 1.7.1 (Regla de conmutación de premisas)**.** *Sea* $\Gamma \cup \{\alpha, \beta, \gamma\}$ *un conjunto de proposiciones, si* $\Gamma \vdash \alpha \to (\beta \to \gamma)$*, entonces* $\Gamma \vdash \beta \to (\alpha \to \gamma)$*.*

Proposición 1.8 (Ley de la doble negación)**.** *Sea* α *una proposición:*

$$\vdash \neg\neg\alpha \to \alpha$$

Demostración. Por el Teorema de Herbrand, es equivalente a ver que $\{\neg\neg\alpha\} \vdash \alpha$. Para ello, usamos la regla de la reducción al absurdo clásica, ya que:

1. $\{\neg\neg\alpha, \neg\alpha\} \vdash \neg\neg\alpha$
2. $\{\neg\neg\alpha, \neg\alpha\} \vdash \neg\alpha$

Luego concluimos que $\{\neg\neg\alpha\} \vdash \alpha$. □

Proposición 1.9 (Ley débil de la doble negación)**.** *Sea* α *una proposición:*

$$\vdash \alpha \to \neg\neg\alpha$$

Demostración. Por el Teorema de Herbrand, es equivalente a ver que $\{\alpha\} \vdash \neg\neg\alpha$. Para ello, usamos la regla de la reducción al absurdo clásica, con lo que partimos que $\{\alpha, \neg\neg\neg\alpha\}$ y tenemos que demostrar una proposición y su negación. Para ello:

1. $\neg\neg\neg\alpha \to \neg\alpha$ por la ley de la doble negación
2. $\neg\neg\neg\alpha$ es una hipótesis
3. $\neg\alpha$ por Modus ponens de 1 y 2
4. α es una hipótesis

Concluimos por la regla de la reducción al absurdo que $\{\alpha\} \vdash \neg\neg\alpha$. □

1.3. Teoremas de coherencia y adecuación

A lo largo de este capítulo hemos manejado en los lenguajes proposicionales dos conceptos fundamentales: las tautologías ($\vDash \alpha$), relacionadas con las interpretaciones; y los teorema ($\vdash \alpha$), relacionados con las demostraciones. Las primeras tienen un gran interés en informática y ciencias de la computación, gracias a las consecuencias semánticas que podemos realizar de forma autómatica, tal y como vimos con el Algoritmo de Davis & Putnam. Por otra parte, las segundas tienen un gran interés matemático, por ser la principal herramienta que sustentan todo el conocimiento matemático. Veremos ahora dos teoremas que nos permiten relacionar las tautologías con los teoremas, de gran importancia en los lenguajes de primer orden.

Teorema 1.10 (de coherencia)**.** *Sea* α *una proposición, si* $\vdash \alpha$*, entonces* $\vDash \alpha$*. Es decir, todo teorema es una tautología.*

Demostración. Si α es un teorema, por definición este tendrá una demostración de n pasos $\alpha_1, \ldots, \alpha_n$:

- α_1 será un axioma y anteriormente probamos que todo axioma era una tautología.
- A α_2 le ocurrirá lo mismo.
- α_i para $i \geqslant 3$ podrá ser un axioma, en cuyo caso ya sabemos cómo proceder o resultado de aplicar modus ponens sobre dos pasos anteriores. Sin embargo, anteriormente vimos que si a y $a \to b$ eran tautologías, entonces b era una tautología, por lo que α_i será una tautología.

Finalmente, llegaremos a $\alpha_n = \alpha$, con lo que podemos concluir que α es una tautología. □

Y además la otra implicación también es cierta, aunque su demostración excede los objetivos del curso.

Teorema 1.11 (de adecuación)**.** *Sea α una proposición, si $\vDash \alpha$, entonces $\vdash \alpha$. Es decir, toda tautología es un teorema.*

2. Lógica de Primer Orden

Es necesario introducir ahora lenguajes en los que podamos cuantificar cosas. Como primer ejemplo, si sabemos que "Todo hombre es mortal" y que "Sócrates es un hombre", nos gustaría deducir que, entonces, "Sócrates es mortal". Sin embargo, para esto hemos de poder cuantificar, cosa que no es posible con los lenguajes proposicionales pero sí con los lenguajes de primer orden.

Los lenguajes de primer orden estarán formados por:

- Constantes: $c_1, c_2, \ldots, a, b, c, \ldots$
- Variables: $x_1, x_2, \ldots, x, y, z, \ldots$
- Símbolos de función: $f_1, f_2, \ldots, f, g, h, \ldots$
- Símbolos de relación: $R_1, R_2, \ldots, R, S, T, \ldots$
- Conectivas lógicas: $\vee$, $\wedge$, $\neg$, $\rightarrow$, $\leftrightarrow$
- Cuantificadores: $\forall$, $\exists$

A los conjuntos de todas las constantes, de todas la variables y de todos los símbolos de función los notaremos por $Cons(\mathcal{L}), Var(\mathcal{L}), Fun(\mathcal{L})$, si $\mathcal{L}$ es nuestro lenguaje de primer orden.

Notación. En otros libros o contextos, en vez de denotar a los símbolos de función o variables con una letra que pueda llevar o no superíndice, estos las denotan con un superíndice:

- $f_1^{n_1}, f_2^{n_2}, \ldots$
- $R_1^{m_1}, R_2^{m_2}, \ldots$

En este caso, el superíndice indica la ariedad de la función o relación. Por ejemplo, si consideramos f^3, tenemos un símbolo de función que se aplica a 3 variables.

Definición 2.1 (Término)**.** Un término es:

1. Cualquier constante.
2. Cualquier variable.
3. Si $t_1, t_2, \ldots, t_n$ son términos y f es un símbolo de función $n-$ario, entonces $f(t_1, t_2, \ldots, t_n)$ es un término.

4. No hay más términos que los que se puedan obtener siguiendo una secuencia finita de pasos a partir de las enunciadas.

Al conjunto de todos los términos de nuestro lenguaje $\mathcal{L}$ lo denotamos por $Term(\mathcal{L})$.

Ejemplo.

- $f(x, f(x, y))$ es un término.
- $f(x, f(x))$ no es un término, ya que usamos un mismo símbolo de función, f, para denotar dos objetos: una función unaria y una función binaria.

Definición 2.2 (Fórmulas atómicas)**.** Si $t_1, \ldots, t_n$ son términos y R es un símbolo de relación $n-$ario, entonces $R(t_1, \ldots, t_n)$ es una fórmula atómica (o simplemente, un átomo).

Definición 2.3 (Fórmulas)**.** Son fórmulas:

1. Las fórmulas atómicas.
2. Si φ y ψ son fórmulas, también lo son:

$$\neg\varphi,\ \varphi \wedge \psi,\ \varphi \vee \psi,\ \varphi \to \psi,\ \varphi \leftrightarrow \psi$$

3. Si x es una variable y φ es una fórmula, también lo son: $\forall x\varphi$, $\exists x\varphi$.
4. No hay más fórmulas que las que se puedan obtener siguiendo una secuencia finita de pasos a partir de las enunciadas.

Al conjunto de todas las fórmulas de nuestro lenguaje $\mathcal{L}$ lo denotamos por $Form(\mathcal{L})$.

Definición 2.4. Una ocurrencia de una variable en una fórmula es una aparición de su escritura.

- En la fórmula $\forall x\varphi$, diremos que φ es el radio de acción de $\forall x$.
- En la fórmula $\exists x\varphi$, diremos que φ es el radio de acción de $\exists x$.

Diremos que x se encuentra cuantificada al ver $\forall x$ o $\exists x$.
Diremos que una ocurrencia de una variable x es ligada si aparece cuantificada o en el radio de acción de $\forall x$ o de $\exists x$.
Finalmente, diremos que una variable es libre si no aparece ligada. Si φ es una fórmula en la que las variables $x_1, \ldots, x_n$ aparecen libres, será usual denotar:

$$\varphi(x_1, \ldots, x_n)$$

Que no debe confundirse con un término de una función o relación $n-$aria, ya que φ no es ni un símbolo de función o relación, sino una fórmula.

Ejemplo. En la siguiente fórmula:

$$\forall x(\exists y R(x, y) \to Q(y))$$

- x aparece cuantificada en su primera ocurrencia.

- y aparece cuantificada en su primera ocurrencia.
- x aparece ligada en su segunda ocurrencia.
- y aparece ligada en su segunda ocurrencia.
- y aparece como variable libre en su tercera ocurrencia.

Definición 2.5 (Sentencia)**.** Una sentencia es una fórmula sin ocurrencias de variables libres.

2.1. Semántica

Trataremos de generalizar el concepto de "interpretación", ya visto para lenguajes proposicionales. Para ello, será necesario primero definir los conceptos de "estructura" y de "asignación".

Definición 2.6 (Estructura)**.** Una estructura ε en un lenguaje $\mathcal{L}$ es una cuádrupla

$$\varepsilon = (D, \{c_i^\varepsilon\}_{n\in\mathbb{N}}, \{f_i^\varepsilon\}_{n\in\mathbb{N}}, \{R_i^\varepsilon\}_{n\in\mathbb{N}})$$

de forma que:

- D es un conjunto no vacío al que llamamos universo o dominio.
- A cada constante c_i de $\mathcal{L}$ le corresponde un elemento c_i^ε de D.
- A cada símbolo de función f_i de $\mathcal{L}$ le corresponde una función $f_i^\varepsilon : D^n \to D$.
- A cada símbolo de relación R_i de $\mathcal{L}$ le corresponde una aplicación $R_i^\varepsilon : D^m \to \mathbb{Z}_2$, de forma que $R_i^\varepsilon(c_1^\varepsilon, c_2^\varepsilon) = 1$ si c_1^ε y c_2^ε están relacionados y 0 en caso contrario.

Definición 2.7 (Asignación)**.** Una asignación v en ε es una aplicación $v : Var(\mathcal{L}) \to D$. Dada una asignación v, podremos extenderla a $v' : Term(\mathcal{L}) \to D$ de la forma:

$$v'(t) = \begin{cases} c^\varepsilon & \text{si } t = c \text{ una constante} \\ v(x) & \text{si } t = x \text{ una variable} \\ f^\varepsilon(v'(t_1), \ldots, v'(t_n)) & \text{si } t = f(t_1, \ldots, t_n) \end{cases}$$

Definición 2.8 (Interpretación)**.** Una interpretación es una tupla (ε, v) con ε una estructura y v una asignación que tiene asociada una aplicación[1] $I_\varepsilon^v : Form(\mathcal{L}) \to \mathbb{Z}_2$ que cumple para cualesquiera fórmulas φ y ψ:

1. $I^v(\neg\varphi) = 1 + I^v(\varphi)$.
2. $I^v(\varphi \wedge \psi) = I^v(\varphi)I^v(\psi)$.
3. $I^v(\varphi \vee \psi) = I^v(\varphi) + I^v(\psi) + I^v(\varphi)I^v(\psi)$.
4. $I^v(\varphi \to \psi) = 1 + I^v(\varphi) + I^v(\varphi)I^v(\psi)$.

[1] A la que próximamente denotaremos simplemente como I^v, por simplicidad, entendiendo que la estructura ε viene dada por el contexto.

5. $I^v(\varphi \leftrightarrow \psi) = 1 + I^v(\varphi) + I^v(\psi)$.
6. $I^v(R(t_1, \ldots, t_n)) = R^{\varepsilon}(v(t_1), \ldots, v(t_n))$
   Con R un símbolo de relación $n-$ario y $t_1, \ldots, t_n$ términos.
7. $I^v(\forall x\varphi) = \begin{cases} 1 & \text{si para todo } a \in D,\ I^{v(x|a)}(\varphi) = 1 \\ 0 & \text{en caso contrario} \end{cases}$
8. $I^v(\exists x\varphi) = \begin{cases} 1 & \text{si existe } a \in D \text{ con } I^{v(x|a)}(\varphi) = 1 \\ 0 & \text{en caso contrario} \end{cases}$

Siendo:

$$v(x \mid a)(y) = \begin{cases} v(y) & \text{si } y \neq x \\ a & \text{si } y = x \end{cases}$$

Definición 2.9. Sea $\varphi \in Form(\mathcal{L})$:

- Dada una estructura ε, diremos que φ es válida en ε si $I^v(\varphi) = 1$ para toda asignación v en ε.
- Dada una estructura ε, diremos que φ es satisfacible en ε si $I^v(\varphi) = 1$ para alguna asignación v en ε.
- Diremos que φ es universalmente válida si φ es válida en cualquier estructura.
- Diremos que φ es satisfacible si existe una estructura ε donde φ es satisfacible.
- Diremos que φ es refutable si $\neg\varphi$ es satisfacible.
- Diremos que φ es una contradicción si $\neg\varphi$ es universalmente válida.

Lema 2.1 (de Coincidencia)**.** *Sea $\varphi \in Form(\mathcal{L})$ de forma que $x_1, \ldots, x_n \in Var(\mathcal{L})$ son las variables con ocurrencias libres en φ y sea ε una estructura. Entonces, dada una asignación v en ε:*

$$I^v(\varphi) = I^w(\varphi)$$

para toda asignación w en ε tal que $w(x_i) = v(x_i)$ para todo $i \in \{1, \ldots, n\}$.

Observación. En particular, si φ es una sentencia, entonces $I^v(\varphi)$ no depende de la asignación v. Por tanto, si φ es satisfacible en ε, entonces φ es válida en ε.

Definición 2.10 (Consecuencia lógica)**.** Sea $\Gamma \cup \{\varphi\} \subseteq Form(\mathcal{L})$, diremos que φ es consecuencia lógica de Γ, notado por $\Gamma \vDash \varphi$, si para toda interpretación (ε, v) tal que $I^v(\gamma) = 1$ para toda $\gamma \in \Gamma$, fuerza a que $I^v(\varphi) = 1$.

Teorema 2.2 (de la Deducción)**.** *Sea $\Gamma \cup \{\varphi, \psi\} \subseteq Form(\mathcal{L})$, son equivalentes:*

1. $\Gamma \vDash \varphi \to \psi$.
2. $\Gamma \cup \{\varphi\} \vDash \psi$.

Definición 2.11 (Inconsistencia)**.** Sea $\Gamma \subseteq Form(\mathcal{L})$, diremos que Γ es inconsistente si no existe una interpretación (ε, v) tal que $I^v(\gamma) = 1$ para toda $\gamma \in \Gamma$.

Teorema 2.3. *Sea $\Gamma \cup \{\varphi\} \subseteq Form(\mathcal{L})$, son equivalentes:*

1. $\Gamma \vDash \varphi$.

2. $\Gamma \cup \{\neg\varphi\}$ *es inconsistente.*

Ejemplo. Demostremos las siguientes fórmulas universalmente válidas:

1. $\vDash \forall x\varphi(x) \leftrightarrow \forall y\varphi(y)$ con[2] y libre para x en $\varphi(x)$.

   Dada cualquier interpretación (ε, v), queremos ver que:

   $$I^v(\forall x\varphi(x) \leftrightarrow \forall y\varphi(y)) = 1$$

   Y sabemos que eso es equivalente a ver que:

   $$I^v(\forall x\varphi(x)) = I^v(\forall y\varphi(y))$$

   Que se puede ver a partir de su definición:

   $$\begin{aligned} I^v(\forall x\varphi(x)) &= \begin{cases} 1 & \text{si } I^{v(x|a)}(\varphi(x)) = 1 \text{ para todo } a \in D \\ 0 & \text{en caso contrario} \end{cases} \\ &\overset{(*)}{=} \begin{cases} 1 & \text{si } I^{v(y|a)}(\varphi(y)) = 1 \text{ para todo } a \in D \\ 0 & \text{en caso contrario} \end{cases} \\ &= I^v(\forall y\varphi(y)) \end{aligned}$$

   Donde en $(*)$ debemos tener cuidado y usar que y es libre para x en $\varphi(x)$, ya que x aparecía libre en φ y como y es libre para x en $\varphi(x)$, al hacer la sustitución de x por y no estaremos cambiando variables libres en φ, por lo que a partir del Lema de Coincidencia (Lema **??**), al no cambiar variables libres en φ, no cambiamos su condición de verdad.

2. $\vDash \forall x\varphi(x) \to \varphi(t)$ con t libre para x en $\varphi(x)$.

   Por el Teorema de la Deducción, probar esto es equivalente a ver que:

   $$\{\forall x\varphi(x)\} \vDash \varphi(t)$$

   Por tanto, sea (ε, v) una interpretación de forma que $I^v(\forall x\varphi(x)) = 1$, entonces para todo $a \in D$, se tendrá $I^{v(x|a)}(\varphi(x)) = 1$. Si tomamos $a = v(t)$, entonces tendremos que:

   $$I^{v(x|a)}(\varphi(x)) = I^v(\varphi(t))$$

2.2. Demostraciones

Trataremos ahora de generalizar lo que hicimos ya para la demostraciones en el caso de los lenguajes proposicionales, para que cualquier demostración hecha con lenguajes proposicionales siga siendo válida ahora.

2.2.1. Definición de una demostración

En lugar de dar directamente la definición de demostración, daremos primero los axiomas de nuestro sistema y las reglas de inferencia que usaremos, para posteriormente dar la definición de demostración.

[2]Estamos usando una notación que se introduce en la siguiente sección.

Axiomas

Sobre nuestro lenguaje $\mathcal{L}$ consideraremos los 3 primeros conjuntos de axiomas, que son los que ya teníamos en lenguajes proposicionales:

$$\mathcal{A}_1 = \{\varphi \to (\psi \to \varphi) : \varphi, \psi \in Form(\mathcal{L})\}$$
$$\mathcal{A}_2 = \{(\varphi \to (\psi \to \chi)) \to ((\varphi \to \psi) \to (\varphi \to \chi)) : \varphi, \psi, \chi \in Form(\mathcal{L})\}$$
$$\mathcal{A}_3 = \{(\neg\varphi \to \neg\psi) \to ((\neg\varphi \to \psi) \to \varphi) : \varphi, \psi \in Form(\mathcal{L})\}$$

Ahora, será necesario considerar nuevos axiomas que nos permitan generalizar lo ya visto para lenguajes proposicionales a lenguajes de primer orden. Como el lector puede deducir, estos axiomas tendrán que contener cuantificadores, ya que es el concepto principal que introducimos en los lenguajes de primer orden. Antes de dar el cuarto axioma[3], introduciremos la siguiente notación:

Notación. Sea $\varphi \in Form(\mathcal{L})$ y $x_1, \ldots, x_n \in Var(\mathcal{L})$, al notar:

$$\varphi(x_1, \ldots, x_n)$$

Estamos diciendo que, si $x_1, \ldots, x_n$ son variables que aparecen en φ, entonces tienen todas sus ocurrencias libres en φ.

Notación. Sea $\varphi \in Form(\mathcal{L})$, $x \in Var(\mathcal{L})$ y $t \in Term(\mathcal{L})$, cuando aparezca:

"t libre para x en $\varphi(x)$"

Y notemos $\varphi(t)$, significará que estamos cambiando las ocurrencias libres de x que había en φ por t. Por ser t un término, este puede depender de otras variables, por lo que en este proceso no se permite que variables de t se queden ligadas, sino que deben aparecer libres.

Podemos dar ya el cuarto axioma:

$$\mathcal{A}_4 = \{\forall x\varphi(x) \to \varphi(t) \mid \varphi \in Form(\mathcal{L}), x \in Var(\mathcal{L}), t \text{ libre para } x \text{ en } \varphi(x)\}$$

Observemos que casos particulares interesantes de este axioma son:

- $\forall x\varphi \to \varphi$
- $\forall x\varphi(x) \to \varphi(x)$

Y podemos finalmente dar el quinto axioma[4]:

$\mathcal{A}_5 = \{\forall x(\varphi \to \psi) \to (\varphi \to \forall x\psi) \mid \varphi, \psi \in Form(\mathcal{L}), x \text{ no aparece libre en } \varphi\}$

De esta forma, nuestro conjunto de axiomas vendrá dado por:

$$\mathcal{A} = \mathcal{A}_1 \cup \mathcal{A}_2 \cup \mathcal{A}_3 \cup \mathcal{A}_4 \cup \mathcal{A}_5$$

Notemos que en estos 5 axiomas no aparecen los conectores $\wedge$, $\vee$, $\leftrightarrow$ ni el cuantificador $\exists$. En caso de querer usarlos:

- Los conectores los expresaremos como fórmulas semánticamente equivalentes pero usando $\neg$ y $\to$.
- Usaremos que $\exists x\varphi$ es semánticamente equivalente a $\neg\forall x\neg\varphi$, siendo $\varphi \in Form(\mathcal{L})$.

[3] Algunos autores dividen este axioma en dos, ya que no consideran la notación que vamos a considerar para poder dar este axioma.

[4] Que en aquellos autores que dividen el cuarto axioma en dos, aparece como el sexto.

Reglas de inferencia

Las reglas de inferencia que consideraremos en nuestro sistema serán las siguientes, las cuales tendremos en cuenta a la hora de realizar la definición de lo que será una demostración:

$$\begin{array}{ll} \text{Modus ponens} & \text{Generalización} \\ \varphi \to \psi & \varphi \\ \varphi & \\ \hline \psi & \forall x\varphi \end{array}$$

Definición 2.12 (Demostración)**.** Si consideramos el conjunto de fórmulas $\mathcal{A}$ previamente definido y sea $\Gamma \cup \{\varphi\} \subseteq Form(\mathcal{L})$, una demostración de p a partir de Γ (notado por $\Gamma \vdash p$) es una secuencia de fórmulas $\alpha_1, \alpha_2, \ldots, \alpha_n$ de forma que $\alpha_n = p$ y se verifica para todo i menor o igual que n alguna de las tres condiciones siguientes:

- $\alpha_i \in \mathcal{A} \cup \Gamma$.
- Existen j, k naturales con $j < k < i$ siendo $\alpha_k = \alpha_j \to \alpha_i$ (Modus ponens).
- Existe un natural j con $j < i$ siendo $\alpha_i = \forall x \alpha_j$ (Generalización).

2.2.2. Primeros resultados

Como primer resultado a destacar, como los lenguajes de primer orden generalizan los lenguajes proposicionales, cualquier demostración para los lenguajes proposicionales seguirán siendo válidas para los lenguajes de primer orden.

Teorema 2.4 (de la Deducción)**.** *Sean $\Gamma \cup \{\varphi, \psi\} \subseteq Form(\mathcal{L})$. Si tenemos que $\Gamma \cup \{\varphi\} \vdash \psi$ y en su demostración no usamos la regla de generalización sobre un paso en que haya intervenido φ con una variable libre en φ, entonces:*

$$\Gamma \vdash \varphi \to \psi$$

Observación. Ha sido necesario introducir la condición extra que no teníamos en lenguajes proposicionales ya que vamos a poder demostrar, por ejemplo, $\{A(x)\} \vdash \forall x A(x)$:

1. $A(x)$ es hipótesis.
2. $\forall x A(x)$ por generalización

Sin embargo, $\nvdash A(x) \to \forall x A(x)$, ya si que si consideramos por ejemplo el dominio $D = \mathbb{Z}_2$ y A es "ser igual a 0", de $A(0)$ no podemos concluir que todo elemento de $\mathbb{Z}_2$ sea 0. Esto se debe a que semánticamente:

$$\{A(x)\} \nvDash \forall x A(x)$$

El lector podría sospechar que la regla de generalización carece de sentido o contradice lo enunciado, pero si somos capaces de demostrar algo para un elemento x arbitrario en un cierto conjunto, entonces seremos capaces de afirmar que $\forall x$ en dicho conjunto, tendremos la proposición conseguida para el elemento arbitrario anterior. Esta es la intuición detrás de la regla de generalización.

Proposición 2.5. *Sean* $\Gamma \cup \{\varphi, \psi\} \subseteq Form(\mathcal{L})$*, si tenemos que* $\Gamma \vdash \varphi \to \psi$*, entonces* $\Gamma \cup \{\varphi\} \vdash \psi$*.*

Teorema 2.6 (Regla de reducción al Absurdo)**.** *Sean* $\Gamma \cup \{\varphi, \psi\} \subseteq Form(\mathcal{L})$ *si tenemos que* $\Gamma \cup \{\neg\varphi\} \vdash \psi$ *y* $\Gamma \cup \{\neg\varphi\} \vdash \neg\psi$ *y en esas demostraciones no usamos la regla de generalización sobre un pasa en el que haya intervenido* $\neg\varphi$ *con una variable libre en* $\neg\varphi$*, entonces:*

$$\Gamma \vdash \varphi$$

Ejemplo. Buscamos demostrar $\vdash (\varphi \to \forall x\psi) \to \forall x(\varphi \to \psi)$ con x no libre en φ.

Buscamos demostrar con precaución[5] que:

$$\{\varphi \to \forall x\psi\} \vdash \forall x(\varphi \to \psi)$$

Para ello:

1. $\varphi \to \forall x\psi$ es una hipótesis.
2. $\forall x\psi \to \psi \in \mathcal{A}_4$
3. $\varphi \to \psi$ por silogismo de 1 y 2.
4. $\forall x(\varphi \to \psi)$ generalización de 3.

Como x no está libre en φ, tampoco lo estará en $\varphi \to \forall x\psi$, por lo que en esta demostración no hemos usado la regla de generalización sobre un paso en el que haya intervenido $\varphi \to \forall x\psi$ con una variable libre en la misma, por lo que podremos aplicar el Teorema de la Deducción, obteniendo lo que queríamos probar.

2.3. Teoremas de adecuación y coherencia

Una parte positiva de los lenguajes de primer orden es que a pesar de ser más generales que los proposicionales, seguimos contando con los teoremas de adecuación y de coherencia:

Teorema 2.7 (de coherencia)**.** *Sea* $\varphi \in Form(\mathcal{L})$*, si* $\vdash \varphi$*, entonces* $\vDash \varphi$*.*

Demostración. La demostración es similar a la del Teorema de coherencia para lenguajes proposicionales, pero ahora hemos de tener en cuenta más axiomas, así como la regla de generalización. □

Teorema 2.8 (de consistencia)**.** *Nuestro conjunto de axiomas* $\mathcal{A}$ *junto con las reglas de inferencia es consistente, es decir, no existe* $\varphi \in Form(\mathcal{L})$ *de forma que* $\vdash \varphi$ *y* $\nvdash \varphi$*.*

Teorema 2.9 (de adecuación)**.** *Sea* $\varphi \in Form(\mathcal{L})$*, si* $\vDash \varphi$*, entonces* $\vdash \varphi$*.*

[5]Para poder aplicar luego el Teorema de la Deducción bajo las hipótesis correctas con la limitación extra.

2.4. Sistemas matemáticos

2.4.1. Lenguajes de Primer Orden con Igualdad

Un lenguaje de primer orden con igualdad es un lenguaje de primer orden $\mathcal{L}$ en el que habrá un símbolo de relación destacado A.

Notación. Aceptaremos las siguientes notaciones con el fin de abreviar los enunciados:

1. Si $s, t \in Term(\mathcal{L})$, usaremos con frecuencia:

$$s = t$$

   Para denotar $A(s, t)$.

2. En el caso de tener $\neg(s = t)$, podremos notar: $s \neq t$.

3. Usaremos $\exists_1 x\varphi(x)$ como abreviatura de:

$$\exists x\varphi(x) \wedge \forall y(\varphi(y) \to x = y)$$

La única diferencia con los lenguajes de primer orden corrientes será que tendremos dos conjuntos de axiomas extras:

$$\mathcal{A}_6 = \{\forall x(x = x) \mid x \in Var(\mathcal{L})\}$$

Y para introducir el último axioma, hace falta introducir más notación:

Notación. Sea $\varphi \in Form(\mathcal{L})$, y $x, y \in Var(\mathcal{L})$, si notamos $\varphi(x, y)$ tras notar $\varphi(x, x)$, significará que estamos reemplazando algunas ocurrencias libres de x por y en la fórmula φ.

El último axioma será:

$$\mathcal{A}_7 = \{(x = y) \to (\varphi(x, x) \to \varphi(x, y)) \mid x, y \in Var(\mathcal{L})\}$$

Observación. Notemos que con dos conjuntos de axiomas que hemos añadido, tenemos que "=" es una relación de equivalencia:

- $\mathcal{A}_6$ nos da la relación reflexiva.
- De $\mathcal{A}_7$ deducimos la simétrica y la transitiva.

Sin embargo, "=" es mucho más que eso, ya que de $\mathcal{A}_7$ no solo deducimos esas propiedades, sino muchas más, tal y como vemos en el siguiente ejemplo.

Ejemplo. Demostraremos que:

$$\vdash (x = y) \to (f(t_1, \ldots, x, \ldots, t_n) = f(t_1, \ldots, y, \ldots, t_n))$$

Para ello, bastará demostrar con cuidado que:

$$\{x = y\} \vdash f(t_1, \ldots, x, \ldots, t_n) = f(t_1, \ldots, y, \ldots, t_n)$$

1. $\forall x(x = x) \in \mathcal{A}_6$
2. $\forall x(x = x) \to f(t_1, \ldots, x, \ldots, t_n) = f(t_1, \ldots, x, \ldots, t_n) \in \mathcal{A}_4$
3. $f(t_1, \ldots, x, \ldots, t_n) = f(t_1, \ldots, x, \ldots, t_n)$ por modus ponens de 1 y 2.
4. $(x = y) \to ((f(t_1, \ldots, x, \ldots, t_n) = f(t_1, \ldots, x, \ldots, t_n)) \to$
   $\to (f(t_1, \ldots, x, \ldots, t_n) = f(t_1, \ldots, y, \ldots, t_n))) \in \mathcal{A}_7$
5. $x = y$ es hipótesis.
6. $(f(t_1, \ldots, x, \ldots, t_n) = f(t_1, \ldots, x, \ldots, t_n)) \to (f(t_1, \ldots, x, \ldots, t_n) = f(t_1, \ldots, y, \ldots, t_n))$
   por modus ponens de 4 y 5.
7. $f(t_1, \ldots, x, \ldots, t_n) = f(t_1, \ldots, y, \ldots, t_n)$ por modus ponens de 3 y 6.

Como no hemos usado ningún paso de generalización, podemos aplicar el Teorema de la Deducción, obteniendo lo que queríamos probar.

2.4.2. Aritmética de Primer Orden

En aritmética de primer orden, consideraremos un lenguaje de primer orden $\mathcal{L}$, que tendrá:

- Variables.
- Una sola constante, que denotaremos por 0.
- Tres símbolos de función, que denotaremos por[6] s, $+$ y $\cdot$.
- Un símbolo de relación que denotaremos por $=$.

En aritméticas de primer orden, consideraremos como axiomas $\mathcal{A}_1, \ldots, \mathcal{A}_7$, junto con los siguientes:

$$\mathcal{N}_1 = \{\forall x(s(x) \neq 0) : x \in Var(\mathcal{L})\}$$
$$\mathcal{N}_2 = \{\forall x \forall y(s(x) = s(y) \to x = y) : x, y \in Var(\mathcal{L})\}$$
$$\mathcal{N}_3 = \{\forall x(x + 0 = x) : x \in Var(\mathcal{L})\}$$
$$\mathcal{N}_4 = \{\forall x \forall y(x + s(y) = s(x + y)) : x, y \in Var(\mathcal{L})\}$$
$$\mathcal{N}_5 = \{\forall x(x \cdot 0 = 0) : x \in Var(\mathcal{L})\}$$
$$\mathcal{N}_6 = \{\forall x \forall y(x \cdot s(y) = x \cdot y + x) : x, y \in Var(\mathcal{L})\}$$
$$\mathcal{N}_7 = \{\varphi(0) \to (\forall x(\varphi(x) \to \varphi(s(x))) \to \forall x \varphi(x)) : \varphi(x) \in Form(\mathcal{L})\}$$

La aritmética de primer orden funcionará bien cuando pensemos que estamos en un dominio similar a $\mathbb{N}$. Pensando esto, los axiomas $\mathcal{N}_i$ conviene entenderlos como:

1. $\mathcal{N}_1$ y $\mathcal{N}_2$ definen cómo funciona la función s, que podemos entender por "siguiente".

[6] Y entenderemos que $\cdot$ tiene mayor prioridad de $+$.

2. $\mathcal{N}_3$ y $\mathcal{N}_4$ definen de forma inductiva la operación $+$.

3. $\mathcal{N}_5$ y $\mathcal{N}_6$ definen de forma inductiva la operación $\cdot$.

4. $\mathcal{N}_7$ es una versión más débil del principio de inducción.

Decimos que $\mathcal{N}_7$ es una versión más débil del principio de inducción porque el principio de inducción (en matemáticas) es el siguiente:

Proposición 2.10. *Sea $A \subseteq \mathbb{N}$, si $0 \in A$ y siempre que $n \in A \implies n+1 \in A$, entonces $A = \mathbb{N}$.*

Donde hacemos una afirmación sobre cualquier subconjunto de $\mathbb{N}$, por lo que estamos considerando elementos dentro de un conjunto no numerable de elementos (el conjunto de todos los subconjuntos de $\mathbb{N}$, que no es numerable). Esta idea no se puede expresar en lenguajes de primer orden, por tener solo un conjunto numerable de fórmulas.

2.4.3. Teoría de conjuntos (de Zermelo-Fraenkel)

La teoría de conjuntos es un lenguaje de primer orden donde:

- No tenemos constantes ni símbolos de función, por lo que los únicos términos que podremos considerar son las variables.
- Como relaciones solo tendremos dos, que denotaremos por $\in$ y $=$ (y usaremos $\notin$ y $\neq$ como sus respectivas negaciones).

Notación. Dados $t, s \in Term(\mathcal{L})$, usaremos $t \subseteq s$ como abreviatura de:

$$\forall x(x \in t \to x \in s)$$

En este contexto, consideraremos como axiomas $\mathcal{A}_1, \ldots, \mathcal{A}_7$, junto con los siguientes (entendiendo que todo lo que sale son variables):

$$ZF_1 = \{x = y \leftrightarrow \forall z(z \in x \leftrightarrow z \in y)\}$$
$$ZF_2 = \{\exists x \forall y(y \notin x)\}$$
$$ZF_3 = \{\forall x \forall y \exists z \forall u(u \in z \leftrightarrow (u = x \vee u = y))\}$$
$$ZF_4 = \{\forall x \exists y \forall z(z \in y \leftrightarrow \exists u(u \in x \wedge z \in u))\}$$
$$ZF_5 = \{\forall x \exists y \forall z(z \in y \leftrightarrow z \subseteq x)\}$$
$$ZF_6 = \{\forall x_1 \exists_1 x_2 \varphi(x_1, x_2) \to \forall x_3 \exists x_4 \forall x_5(x_5 \in x_4 \leftrightarrow \exists x_6(x_6 \in x_3 \wedge \varphi(x_6, x_5)))\}$$
$$ZF_7 = \{\exists x(\emptyset \in x \wedge \forall y(y \in x \to y \cup \{y\} \in x))\}$$
$$ZF_8 = \{\forall x(x \neq \emptyset \to \exists y(y \in x \wedge \neg \exists z(z \in x \wedge z \in y)))\}$$

Las variables recibirán usualmente el nombre de "conjuntos" o "elementos" y una forma de entender mejor estos axiomas es la siguiente:

1. ZF_1 recibe el nombre de "extensionalidad" y puede entenderse como una condición de cuándo dos conjuntos son iguales.

2. ZF_2 afirma la existencia de un conjunto sin elementos.

   A partir de ZF_1 y ZF_2 puede demostrarse que aquel conjunto sin elementos es único. Por tanto, a partir de ahora nos referiremos a este único conjunto por $\emptyset$, y le llamaremos "conjunto vacío". De esta forma, ZF_2 recibe el nombre de "existencia del conjunto vacío".

3. ZF_3 recibe el nombre de "emparejamiento", y afirma que dados dos conjuntos x e y, podemos considerar z, el conjunto formado por estos dos elementos: $\{x, y\}$. En el caso $\{x, x\}$, notaremos simplemente $\{x\}$.

4. ZF_4 nos dice que siempre que tengamos un conjunto x, existirá un conjunto y que contendrá todos aquellos elementos que están en algún conjunto de x. De esta forma, podemos pensar en ZF_4 como en la existencia de las uniones arbitrarias de conjuntos. Si consideramos la unión de un conjunto x, podremos denotarlo por:

   $$\bigcup x$$

   Y cuando tengamos dos conjuntos t y s, podremos denotar:

   $$t \cup s = \bigcup \{t, s\}$$

5. ZF_5 recibe el nombre de "conjunto potencia" y afirma que dado un conjunto x, existe un conjunto y que contiene todos aquellos conjuntos que sean subconjuntos de x. Este conjunto y recibirá usualmente el nombre de $\mathcal{P}(x)$.

6. ZF_6 recibe el nombre de "esquema de reemplazo" y nos permite definir funciones mediante la regla $\varphi(x_1, x_2)$: dado cualquier x_1, existirá un único x_2 de forma que se tenga $\varphi(x_1, x_2)$. En cuyo caso, podremos considerar el conjunto imagen de un conjunto por dicha aplicación:

   Si pensamos en x_3 como un subconjunto del dominio de la aplicación, entonces existirá un x_4 (imagen de x_3 por la aplicación), y este verificará que un elemento está en él si y solo si hay un elemento de x_3 que se aplicaba en él.

7. ZF_7 recibe el nombre de "axioma del infinito", y es que afirma la existencia de un conjunto x que contiene a $\emptyset$ y es infinito.

8. ZF_8 recibe el nombre de "axioma de regularidad", y viene a decir que dada conjunto no vacío x contiene un elemento y que es disjunto con el propio x (es decir, que cualquier conjunto no vacío contiene un elemento que no comparte ningún elemento con el propio conjunto no vacío de partida).

   Este axioma nos permite no caer en cadenas infinitas de pertenencias. Por ejemplo, si tuviéramos dos conjuntos x e y de forma que:

   $$x = \{y\} \qquad y = \{x\}$$

   Entonces, tendríamos:

   $$x \ni y \ni x \ni y \ni \dots$$

   Pero este axioma no lo permite.

Axioma de elección (AE).
Para todo conjunto no vacío x existe un conjunto y que tiene un único elemento en común con cada elemento de x.

Lema de Zorn.
Si toda cadena de un conjunto ordenado tiene cota superior, entonces el conjunto tiene un elemento maximal.

Principio de buena ordenación.
Todo conjunto no vacío admite un buen orden (un elemento mínimo).

Hipótesis del continuo (HC).
Todo subconjunto infinito de $\mathbb{R}$ es numerable o tiene la misma cardinalidad de que $\mathbb{R}$.

(AE) y (HC) no son demostrables con la axiomática de Zermelo-Fraenkel, son independientes entre sí y son consistentes con estos axiomas, así como sus negaciones.

3. Introducción a la Teoría Descriptiva de Conjuntos

La Teoría Descriptiva de Conjuntos (TDC a partir de ahora) tiene como un objetivo clasificar enunciados o fórmulas según su complejidad, concepto que luego formalizaremos. Por ejemplo, si consideramos sobre las funciones del intervalo $[0,1]$ en $\mathbb{R}$ la propiedad de "ser continua":

$$\forall x \in [0,1]\ \forall \varepsilon > 0\ \exists \delta > 0 : \forall x_0\ |x - x_0| < \delta \Longrightarrow |f(x) - f(x_0)| < \varepsilon$$

Como sabemos que toda función continua en $[0,1]$ es uniformemente continua y que toda función uniformemente continua es continua, podemos reescribir esta propiedad de una forma "menos compleja", eliminando en ε la dependencia de x

$$\forall \varepsilon > 0\ \exists \delta > 0 : \forall x_0\ |x - x_0| < \delta \Longrightarrow |f(x) - f(x_0)| < \varepsilon$$

y obteniendo así una fórmula más corta, algo que nos interesará, puesto que intentaremos buscar las fórmulas más cortas (en función de las variables y cuantificadores que aparecen en ellas) que nos definan ciertas propiedades.

3.1. Construcción de fórmulas

Sin olvidar que la teoría sobre la que trabajamos (la de Zermelo-Fraenkel) es en particular un lenguaje de primer orden, recordamos que nuestras fórmulas van a estar formadas por, fijado un conjunto X que contendrá los elementos a los que nos refiramos:

- Términos, referidos a objetos.
- Conectores: $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$.
- Cuantificadores: $\forall$, $\exists$.

De esta forma, una fórmula para nosotros será una composición finita de estos elementos.

Estaremos especialmente interesados en el hecho de que las fórmulas sean composiciones finitas de dichos elementos, así como en estudiar los cuantificadores que aparecen en las fórmulas.

Ejemplo. Si consideramos $X = \mathbb{N}$ y en este contexto consideramos la aplicación "sucesor" $s : \mathbb{N} \to \mathbb{N}$, la fórmula:

$$s \leqslant s(n)$$

Contiene a n como variable libre, y (en general) podrá ser cierta o falsa en función de las sustituciones de n realicemos (aunque en este caso hemos considerado una fórmula que ante cualquier sustitución siempre es cierta).

A partir de ahora, lo que nos interesará es dada una fórmula P en la que hay una variable libre y trabajando sobre un conjunto X, podremos siempre considerar por el axioma de composición el conjunto formado por aquellos elementos de X para los cuales la fórmula P sea cierta:

$$X_P = \{x \in X \mid P(x)\}$$

Sin embargo, el recíproco de esta afirmación (que para cada cojunto siempre podemos encontrar una fórmula que cumplan exclusivamente los elementos del conjunto) no será generalmente cierta. Por ejemplo, si consideramos X un conjunto finito, como $\mathcal{P}(X)$ es finito, siempre podremos hacerlo; pero en un caso general con X cualquier conjunto (posiblemente no numerable), podemos pensar en que la cantidad de fórmulas que podemos construir es un conjunto numberable, por lo que no llegaremos a abarcar todas las posibilidades[1].

Ejemplo. Como primeros ejemplos de conjuntos a destacar, fijado un conjunto X, consideramos una cantidad numerable de subconjuntos de X: $\{X_n\}_{n\in\mathbb{N}}$ con $X_n \subseteq X$ para todo $n \in \mathbb{N}$.

- Si pensamos en la intersección de todos estos conjuntos:

  $$\bigcap_{n\in\mathbb{N}} X_n$$

  Podemos tratar de buscar una fórmula que defina única y exclusivamente a todos los elementos de este conjunto, como por ejemplo:

  $$\forall n(n \in \mathbb{N} \Longrightarrow x \in X_n)$$

- Si ahora consideramos la unión de todos ellos:

  $$\bigcup_{n\in\mathbb{N}} X_n$$

  Y tratamos de buscar una fórmula que lo defina, llegamos a:

  $$\exists n(n \in \mathbb{N} \wedge x \in X_n)$$

A partir de este primer ejemplo y viendo la relación existente entre los cuantificadores $\forall$ y $\exists$ con las operaciones $\cap$ y $\cup$, buscamos ahora cómo podemos expresar ciertas fórmulas sobre algún espacio X previamente fijado en forma de conjuntos, proceso que ilustraremos con los siguientes ejemplos.

[1] Es mucho más complejo que esto, este argumento no es suficiente para demostrarlo.

Ejemplo. En cada caso, consideraremos un conjunto X distinto.

- En el espacio de las sucesiones de números reales: $X = \mathbb{R}^{\mathbb{N}} = \{x : \mathbb{N} \to \mathbb{R}\}$

  Pensamos en la propiedad de que una sucesión sea "casi nula", que intuitivamente podemos definir como que una sucesión tenga una cantidad infinita de términos nulos. De manera formal, podemos escribir que existe un término a partir del cual todos los términos de la sucesión son cero:

  $$\exists n \in \mathbb{N}\ \forall m \geqslant n\ x(m) = 0$$

  Que podemos escribir de forma más rigurosa como (entendiendo que donde pone $m \geqslant n$ deberíamos escribir $m \geqslant n \wedge m \in \mathbb{N}$):

  $$\exists n(n \in \mathbb{N} \wedge \forall m(m \geqslant n \Longrightarrow x(m) = 0))$$

  Donde observamos que en esta la variable x aparece libre, por lo que podemos tratar de buscar un conjunto que contenga todos aquellos elementos que cumplan la fórmula para cierto x y ninguno más, conjunto al que denotaremos por C_{00}.

  Para hayar este conjunto, lo que haremos será en primer lugar considerar los términos que aparecen en la fórmula y en segundo lugar, tratar de relacionarlos con los conectores y cuantificadores que aparecen. Para ello, los términos que aparecen en la fórmula son:

  $$n \in \mathbb{N} \qquad m \geqslant n \qquad x(m) = 0$$

  Y para construir el conjunto que venga definido por la fórmula, lo que haremos será ir poco a poco de dentro hacia afuera, considerando primero el conjunto que cumpla:

  $$x(m) = 0$$

  De esta forma, fijado m, definimos:

  $$X_m = \{x \in \mathbb{R}^{\mathbb{N}} : x(m) = 0\}$$

  Ahora, buscamos el conjunto que venga definido por la fórmula:

  $$\forall m(m \geqslant n \Longrightarrow x(m) = 0)$$

  Que podemos reescribir como:

  $$\forall m(m \geqslant n \Longrightarrow x \in X_m)$$

  Observando el $\forall$, podemos pensar en reescribir este conjunto como en una intersección de conjuntos:

  $$\bigcap_{m \geqslant n} X_m$$

  Finalmente, la fórmula entera:

  $$\exists n(n \in \mathbb{N} \wedge \forall m(m \geqslant n \Longrightarrow x(m) = 0))$$

La podemos reescribir como:

$$\exists n \left(n \in \mathbb{N} \wedge x \in \bigcap_{m \geqslant n} X_m \right)$$

Que podemos expresar ahora como la unión de ciertos conjunto.

$$\bigcup_{n \in \mathbb{N}} \bigcap_{m \geqslant n} X_m$$

Obteniendo así nuestro conjunto C_{00}:

$$C_{00} = \bigcup_{n \in \mathbb{N}} \bigcap_{m \geqslant n} X_m$$

- Sobre el mismo espacio $X = \mathbb{R}^{\mathbb{N}}$ podemos ahora considerar la propiedad de "ser convergente a 0", propiedad definida por:

$$\forall \varepsilon > 0 \; \exists m \in \mathbb{N} \; \forall n \geqslant m \Longrightarrow |x(m)| < \varepsilon \tag{3.1}$$

Y que de forma rigurosa puede escribirse como (entendiendo que donde pone $n \geqslant m$ deberíamos escribir $n \geqslant m \wedge n \in \mathbb{N}$ y que donde pone $\varepsilon > 0$ deberíamos poner $\varepsilon \in \mathbb{R}^+$):

$$\forall \varepsilon (\varepsilon > 0 \Longrightarrow \exists m (m \in \mathbb{N} \wedge \forall n (n \geqslant m \Longrightarrow |x(m)| < \varepsilon)))$$

Fórmula a partir de la cual podemos extrar un conjunto de igual forma que hicimos anteriormente, identificando que los términos son:

$$\varepsilon > 0 \qquad m \in \mathbb{N} \qquad n \geqslant m \qquad |x(m)| < \varepsilon$$

Y construyendo la fórmula de dentro hacia afuera. Para ello, en primer lugar, fijados m y ε definimos el conjunto:

$$X_{n,\varepsilon} = \{x \in \mathbb{R}^{\mathbb{N}} : |x(n)| < \varepsilon\}$$

Y posteriormente escribiendo las sucesivas fórmulas como uniones e intersecciones, llegando a:

$$\bigcap_{\varepsilon > 0} \bigcup_{m \in \mathbb{N}} \bigcap_{n \geqslant m} X_{n,\varepsilon}$$

Sin embargo, hay una diferencia ahora entre la fórmula obtenida anteriormente y esta; resulta que en esta fórmula estamos considerando una intersección no numerable de elementos, al considerar la intersección de todos aquellos elementos de $\mathbb{R}^+$, un hecho que nos va a dificultar luego algo en lo que estamos interesados[2].

A pesar de ello, la solución en este caso es bien sencilla. Resulta que la definición de convergencia a 0 cuya definición escribimos en (**??**) puede caracterizarse

[2]Que es poder definir una sigma álgebra que contenga uniones e intersecciones numerables de ciertos conjuntos.

en función de una sucesión convergente a cero, pudiendo cambiar la fórmula que describe la propiedad de "ser convergente a cero" por la fórmula:

$$\forall k \in \mathbb{N}\ \exists m \in \mathbb{N}\ \forall n \geqslant m \Longrightarrow |x(m)| < \frac{1}{k}$$

Si ahora realizamos nuevamente el proceso anterior de reescribir a qué conjunto llegamos, obtenemos ahora sí un conjunto dado por intersecciones y uniones numerables de ciertos conjuntos:

$$\bigcap_{k\in\mathbb{N}} \bigcup_{m\in\mathbb{N}} \bigcap_{n\geqslant m} X_{n,\frac{1}{k}}$$

- Si ahora consideramos $X = \{f : \mathbb{R} \to \mathbb{R} : f \text{ continua}\}$ y fijado $L \in \mathbb{R}^+$, pensamos en la propiedad de $\lim\limits_{x\to\infty} f(x) = L$, definida por:

$$\forall \varepsilon > 0\ \exists M > 0\ \forall x > M\ |f(x) - L| < \varepsilon$$

Se nos plantea el problema anterior de que obtendríamos uniones o intersecciones no numerables de conjuntos. Sin embargo, vemos que es fácil reemplazar ε y M para que esto no suceda, obteniendo una expresión equivalente:

$$\forall k \in \mathbb{N}\ \exists M \in \mathbb{N}\ \forall x > M\ |f(x) - L| < \frac{1}{k}$$

Sin embargo, seguimos teniendo el problema de que $x \in \mathbb{R}$, que pensamos en cómo solucionar. Como estamos considerando funciones continuas de $\mathbb{R}$ en $\mathbb{R}$ y $\mathbb{Q} \subseteq \mathbb{R}$ es denso y numerable, podemos considerar $x \in \mathbb{Q}$ y reescribir la propiedad de forma equivalente como:

$$\forall k \in \mathbb{N}\ \exists M \in \mathbb{N}\ \forall x \in \mathbb{Q}, x > M\ |f(x) - L| < \frac{1}{k}$$

Así, obtenemos una fórmula:

$$\forall k \left(k \in \mathbb{N} \Longrightarrow \exists M \left(M \in \mathbb{N} \wedge \forall x \left(x \in Q \wedge x > M \Longrightarrow |f(x) - L| < \frac{1}{k} \right) \right) \right)$$

Fijado x y k, definimos el conjunto:

$$X_{x,k} = \left\{ f \in C(\mathbb{R}) : |f(x) - L| < \frac{1}{k} \right\}$$

Y la fórmula nos da el conjunto:

$$\bigcap_{k\in\mathbb{N}} \bigcup_{M\in\mathbb{N}} \bigcap_{\substack{x>M \\ x\in\mathbb{Q}}} X_{x,k}$$

- Finalmente, si ahora consideramos el mismo espacio X y la pensamos en la propiedad de que una función tenga límite en infinito:

$$\forall L \in \mathbb{R}\ \forall \varepsilon > 0\ \exists M > 0\ \forall x > M\ |f(x) - L| < \varepsilon$$

Sabemos que podemos sustituir ε, M y x para obtener intersecciones y uniones numerables, pero ¿cómo podemos ahora hacer esto con L? Pues bien, podemos usar un resultado bien conocido, y es que $\mathbb{R}$ es completo, por lo que cualquier sucesión de Cauchy es convergente y viceversa, con lo que podemos reescribir esta fórmula en una equivalente de la forma:

$$\forall k \in \mathbb{N}\ \exists M \in \mathbb{N}\ \forall x, y \in \mathbb{Q}, x, y > M\ |f(x) - f(y)| < \varepsilon$$

Con esta gran cantidad de ejemplos hemos visto cómo podemos obtener conjuntos a partir de fórmulas, así como tratar de buscar siempre una cantidad numerable de intersecciones y uniones, que generalmente obtendremos usando teoremas fundamentales del espacio en el que trabajemos.

En resumen, fijado un conjunto X, nos interesará dar propiedades mediante fórmulas, a partir de las cuales construir conjuntos mediante intersecciones y uniones (preferiblemente numerables) de conjuntos, las cuales vendrán dadas por los cuantificadores que hemos usado en la fórmula para describir una propiedad específica. De esta forma, dada una cierta propiedad y considerando una cierta topología $\mathcal{T}$ sobre el espacio X, los conjuntos que obtengamos podrán ser:

- Abiertos.
- Cerrados.
- Intersecciones numerables de abiertos.
- Uniones arbitrarias de cerrados.
- Uniones numerables de intersecciones numerables de abiertos.
- Intersecciones numerables de uniones numerables de cerrados.
- ...

De esta forma, llegaremos luego a considerar una $\sigma-$álgebra de Borel, que será donde podamos trabajar.

Finalmente nos preguntamos si toda fórmula puede reducirse a unos cuantificadores numerables, pregunta cuya respuesta será que no[3], y esta respuesta nos hará interesarnos por una noción más general de los cardinales de los conjuntos.

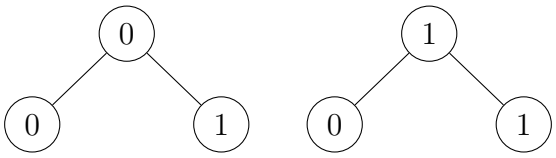
3.2. Conjunto de Cantor

Definición 3.1 (Conjunto de Cantor)**.** Definimos el conjunto de Cantor como el conjunto de las sucesiones de $\{0, 1\}$:

$$2^{\mathbb{N}} = \{f : \mathbb{N} \to \{0, 1\}\}$$

[3]Se verá un leve razonamiento de por qué.

Las sucesiones de números de $\{0,1\}$ podemos entenderlas de forma gráfica como cada una de las elecciones (arriba o abajo) en las bifurcaciones del siguiente dibujo:



Notación. Nos interesará también considerar una sucesión finita de elementos de $\{0,1\}$:

$$2^{<\mathbb{N}} = \bigcup_{n\in\mathbb{N}} \{f : \{0,1,\dots,n\} \to \{0,1\}\}$$

Lo que nos interesará será comparar qué tan cercanas están los elementos del conjunto de Cantor entre sí de una forma intuitiva.

Ejemplo. Por ejemplo, ante las sucesiones:

$$(0,1,0,1,0,\dots)$$
$$(0,1,1,0,0,\dots)$$
$$(0,1,0,1,1,\dots)$$

Nos interesará decir que la tercera sucesión es la que más se parece a la primera. Una analogía es comparar lo que estamos haciendo con los números irracionales, ya que el conjnto de los números irracionales, $\mathbb{R}\setminus\mathbb{Q}$, podemos verlo como un entero seguido de una sucesión de naturales del conjunto $\{0,1,\dots,9\}$:

$$x \in \mathbb{R}\setminus\mathbb{Q}, \quad x = a' n_1 n_2 n_3 \dots \qquad a \in \mathbb{Z}, n_1, n_2, n_3, \dots \in \{0,1,\dots,9\}$$

Ante los siguientes números irracionales:

$$3{,}141529\dots$$
$$3{,}141578\dots$$
$$3{,}142386\dots$$

Decimos que el segundo es más cercano al primero que el tercero al primero.

Definición 3.2. Definimos en el conjunto de Cantor la distancia: $d : 2^{\mathbb{N}} \times 2^{\mathbb{N}} \to \mathbb{R}$ dada por:

$$d(x,y) = \begin{cases} \frac{1}{2^{n+1}} & \text{si } x \neq y, \quad n = \min\{k \in \mathbb{N} \mid x(k) \neq y(k)\} \\ 0 & \text{en otro caso} \end{cases}$$

En el conjunto $2^{<\mathbb{N}}$, fijado $s \in 2^{<\mathbb{N}}$ podemos definir el siguiente conjunto de abiertos básicos para s (siendo n la longitud de s):

$$O_s = \{x \in 2^{\mathbb{N}} \mid x_{|\{0,1,\dots,n\}} = s\} = \left\{x \in 2^{\mathbb{N}} \mid d(s,x) \leqslant \frac{1}{2^{n+1}}\right\}$$

Ejercicio 3.2.1. Demostrar que $(2^{\mathbb{N}}, d)$ es un espacio completo.

Recordemos que un espacio métrico (X, d) es completo si toda sucesión de Cauchy es convergente.

Definición 3.3 (Polaco)**.** Sea (X, d) un espacio métrico, decimos que es polaco si admite una distancia que lo hace completo e induce un espacio topológico separable.

Ejemplo. Un ejemplo de un espacio métrico polaco es el conjunto de los irracionales con la distancia usual de $\mathbb{R}$ inducida, ya que el espacio métrico que consideramos no es completo (hay sucesiones de irracionales que convergen a racionales), pero con la distancia análoga a la anterior definida sí que es completo.

Lema 3.1 (de Cantor)**.** *Todo espacio (no numerable) completo*[4] *separable y sin puntos aislados admite un conjunto homeomorfo al conjunto de Cantor.*

Demostración. Sea X un espacio bajo esas hipótesis, pensamos la demostración como si estuviéramos en $[0, 1]$:

1. Cogemos dos puntos suficientemente lejos, x_0 y x_1, de forma que podamos coger $\varepsilon_0, \varepsilon_1 \in \mathbb{R}^+$ de forma que[5]:
$$B(x_0, \varepsilon_0) \cap B(x_1, \varepsilon_1) = \emptyset$$
2. Dentro de $B(x_0, \varepsilon_0)$ volvemos a coger dos puntos en estas condiciones: x_{00} y x_{01}, de forma que:
$$B(x_{00}, \varepsilon_{00}) \cap B(x_{01}, \varepsilon_{01}) = \emptyset$$
Repetimos el procedimiento en $B(x_1, \varepsilon_1)$ y hacemos una inducción.
3. De esta forma, por cada punto de $2^{\mathbb{N}}$ hemos encontrado una sucesión de bolas abiertas decrecientes. Se verifica que la intersección de todas ellas es un único punto[6]:
$$\bigcap B_x = \{p\} \qquad p \in X$$
Hemos conseguido una aplicación $\Phi : 2^{\mathbb{N}} \to X$, entre el conjunto de Cantor y el nuestro.
4. Demostramos que:
   - Φ es inyectiva.
   - $2^{\mathbb{N}}$ es compacto.
   - Φ es continua[7].

   Con lo que, como X es Hausdorff, una aplicación continua de un cerrado en Hausdorff es cerrada, con lo que acabamos deduciendo que Φ es homeomorfismo sobre su imagen.

□

Ejercicio 3.2.2. Formalizar la demostración anterior.

[4] En realidad, bastaría con ser polaco.
[5] Usar que no hay puntos aislados.
[6] Usar la completitud.
[7] Usar que es separable.

3.3. Jerarquía de Borel

Fijado un espacio topológico procedente de una métrica, $(X, \mathcal{T})$, definimos:

$$\Sigma_0 = \{U \subseteq X \mid U \text{ abierto}\} \qquad \Sigma_1 = \left\{ \bigcup_{n\in\mathbb{N}} A_n \mid \text{ cerrado} \right\}$$

$$\Pi_0 = \{U \subseteq X \mid U \text{ cerrado}\} \qquad \Pi_1 = \left\{ \bigcap_{n\in\mathbb{N}} \mid A_n \text{ abierto} \right\}$$

Ejercicio 3.3.1. Sea $x \in X$ y $A \subseteq X$, definimos la distancia entre a y X como:

$$d(x, A) = \text{ínf}\{d(x, a) \mid a \in A\}$$

Dado $r > 0$, se verifica que el conjunto:

$$\{x \in X \mid d(x, A) < r\}$$

es un abierto.

Proposición 3.2. *Se verifica que:*

1. $\Sigma_0 \subseteq \Pi_1$.
2. $\Pi_0 \subseteq \Sigma_1$.
3. $\Pi_0 \subseteq \Pi_1$.
4. $\Sigma_0 \subseteq \Sigma_1$.

Demostración. Vemos las inclusiones:

1. Sea $A \in \Sigma_0$:
$$A = \bigcup_{n\in\mathbb{N}} A$$
2. Sea $A \in \Pi_0$:
$$A = \bigcap_{n\in\mathbb{N}} A$$
3. Sea $C \in \Pi_0$, veamos que C puede escribirse como una intersección numerable de abiertos. Para ello, sabemos que:
$$C = \overline{C} = \{x \in X \mid d(x, C) = 0\}$$
Por lo que podemos tomar:
$$C = \bigcap_{n\in\mathbb{N}} \left\{ x \in X \mid d(x, C) < \frac{1}{n} \right\}$$

4. Sea $C \in \Sigma_0$, sabemos que $X \setminus C \in \Pi_0$, por lo que podemos escribir $X \setminus C$ como intersección numerable de abiertos:

$$X \setminus C = \bigcap_{n \in \mathbb{N}} A_n$$

Tomando complementario:

$$C = X \setminus (X \setminus C) = X \setminus \left(\bigcap_{n \in \mathbb{N}} A_n \right) = \bigcup_{n \in \mathbb{N}} X \setminus A_n$$

Tenemos que $X \setminus A_n$ es un cerrado, $\forall n \in \mathbb{N}$, luego $C \in \Sigma_1$.

□

Si seguimos definiendo conjuntos para la jerarquía:

$$\Sigma_1 \qquad \Sigma_2 = \left\{ \bigcup_{n \in \mathbb{N}} A_n \mid A_n \in \Pi_j, j \in \{0,1\} \right\}$$

$$\Pi_1 \qquad \Pi_2 = \left\{ \bigcap_{n \in \mathbb{N}} A_n \mid A_n \in \Sigma_j, j \in \{0,1\} \right\}$$

Y por inducción, dado $n \in \mathbb{N}$, definimos:

$$\Sigma_{n-1} \qquad \Sigma_n = \left\{ \bigcup_{n \in \mathbb{N}} A_n \mid A_n \in \Pi_j, j < n \right\}$$

$$\Pi_{n-1} \qquad \Pi_n = \left\{ \bigcap_{n \in \mathbb{N}} A_n \mid A_n \in \Sigma_j, j < n \right\}$$

Proposición 3.3. *Se verifica que:*

1. $\Sigma_j \subseteq \Sigma_n \ \forall j < n$.

2. $\Pi_j \subseteq \Pi_n \ \forall j < n$.

Ejercicio 3.3.2. Hacer la Proposición anterior.

Notación. Si previamente fijamos un espacio X sobre el que trabajar, escribiremos:

$$\Sigma_n(X) \qquad \Pi_n(X) \qquad \forall n \in \mathbb{N}$$

De esta forma, podemos también hablar sobre estos conceptos sin fijar antes ningún conjunto, en cuyo caso no serán conjuntos, sino clases, y notaremos:

$$\Sigma_n \qquad \Pi_n \qquad \forall n \in \mathbb{N}$$

3.3.1. Conjunto de Vitali

Un problema que surgió de forma natural en matemáticas cuando se intentó definir una medida, es decir, una función $\mu : \cdot \to \mathbb{R}_0^+$ que verifique unas ciertas propiedades deseables con el fin de obtener una medida del tamaño de cualquier conjunto, propiedades como:

- Que la medida de los intervalos sea la diferencia de los extremos:

$$\mu([a, b]) = b - a$$

- La $\sigma-$aditividad de la medida:

$$\mu\left(\biguplus_{n\in\mathbb{N}} A_n\right) = \sum_{n=0}^{\infty} \mu(A_n)$$

- La medida es invariante frente a traslaciones:

$$\mu(t(A)) = \mu(A)$$

De estas tres propiedades podía deducirse que si tenemos $A \subseteq B \subseteq C$, entonces:

$$\mu(A) \leqslant \mu(B) \leqslant \mu(C)$$

Sin embargo, no podemos considerar esta medida sobre todos los conjuntos a considerar, porque tendríamos entonces varias contradicciones en la teoría, tal y como veremos a continuación. Por esta razón, las medidas (como por ejemplo la de Lebesgue, ya estudiada en Análisis Matemático II) se definen sobre ciertos conjuntos restringidos de $\mathcal{P}(X)$, no sobre todo este conjunto. En el ejemplo de la medida de Lebesgue, estos conjuntos eran los medibles, $\mathcal{M}$. En este caso, los conjuntos borelianos (de la $\sigma-$álgebra de Borel) estaban dentro de los medibles.

En teorías matemáticas constructivas se verifica que todos los conjuntos que se consideran son mediables, pero al considerar axiomas o resultados como el Axioma de Elección, empiezan a surgir ejemplos de conjuntos no medibles.

En nuestro caso, la jerarquía definida hasta el momento, el conjunto de todos los Σ_n y Π_n no nos es suficiente. Por ejemplo, la unión de todos los conjuntos Σ_n está dentro del $\sigma-$álgebra de Borel, que ni siquiera llega a todos los conjuntos medibles, por lo que queremos extender nuestra jerarquía con el fin de llegar a poder clasificar más tipos de conjuntos.

Lo que haremos en esta sección será dar un ejemplo de un conjunto no medible (es decir, un conjunto que pone en conflicto alguna de las propiedades enunciadas que son buenas para las medidas). Como este conjunto no será medible, en particular no será de Borel, mucho menos puede estar en nuestra jerarquía.

Definición 3.4 (Conjunto de Vitali)**.** Sea $V \subseteq \mathbb{R}$, decimos que V es un conjunto de Vitali si:

$$V \cap \{x + \mathbb{Q}\} = \{y\} \subseteq \mathbb{R} \qquad \forall x \in \mathbb{R}$$

Proposición 3.4. *Todo conjunto de Vitali $V \subseteq [0,1]$ verifica:*

1. $(q+V) \cap (p+V) = \emptyset \qquad \forall p, q \in \mathbb{Q}, p \neq q$
2. $[0,1] \subseteq \bigcup_{\substack{q\in\mathbb{Q}\\ -1\leqslant q\leqslant 1}} (q+V) \subseteq [-1,2]$

Demostración. Supuesto que tenemos un conjunto de Vitali $V \subseteq [0,1]$, veamos las dos propiedades:

1. Supongamos que tenemos $p, q \in \mathbb{Q}$ con $p \neq q$ de forma que:
$$(q+V) \cap (p+V) \neq \emptyset$$
En dicho caso, $\exists u, v \in V$ de forma que:
$$q + u = p + v$$
En cuyo caso, $v = u + (q-p)$, por lo que $v \neq u$. Sin embargo, tenemos entonces que:
$$\{u, v\} \subseteq V \cap \{u + \mathbb{Q}\}$$
Por lo que llegamos a una contradicción con que V era un conjunto de Vitali.

2. Para la primera inclusión, sea $x \in [0,1]$, basta observar que:
$$V \cap (x + \mathbb{Q}) = \{y\} \subseteq \mathbb{R}$$
Para la segunda:
$$\bigcup_{\substack{q\in\mathbb{Q}\\ -1\leqslant q\leqslant 1}} (q+V) \subseteq \bigcup_{\substack{q\in\mathbb{Q}\\ -1\leqslant q\leqslant 1}} (q+[0,1]) \subseteq [-1,2]$$

□

Todavía no demostraremos la existencia del mismo, pero supuesta la existencia y vistas estas propiedades, podemos ver por un lado que aplicando la $\sigma-$aditividad:

$$\mu\left(\bigcup_{\substack{q\in\mathbb{Q}\\ -1\leqslant q\leqslant 1}} (q+V)\right) = \sum_{\substack{q\in\mathbb{Q}\\ -1\leqslant q\leqslant 1}} \mu(q+V) = \sum_{\substack{q\in\mathbb{Q}\\ -1\leqslant q\leqslant 1}} \mu(V)$$

Pero como esta cantidad ha de ser menor que 3, por estar contenido el conjunto en $[-1,2]$, esta última serie a de ser finita, luego ha de ser $\mu(V) = 0$.

Por otra parte, observamos que:

$$1 = \mu([0,1]) \leqslant \mu\left(\bigcup_{\substack{q\in\mathbb{Q}\\ -1\leqslant q\leqslant 1}} (q+V)\right) \leqslant \mu[-1,2] = 3$$

Por lo que no puede ser $\mu(V) = 0$.

Vemos que supuesta la existencia de este conjunto (algo que no hemos demostrado todavía), llegamos a una contradicción, por lo que este conjunto (en caso de existir), no podrá ser medible, y mucho menos estar en nuestra jerarquía.

Demostración. Para probar la existencia de un conjunto de Vitali en $[0,1]$, lo que haremos será definir la siguiente relación de equivalencia:

$$\forall x, y \qquad x \sim y \iff \exists q \in \mathbb{Q} : x = q + y$$

Con esta relación de equivalencia, podemos considerar la proyección al cociente:

$$\mathbb{R} \xrightarrow{\pi} \mathbb{R}/\sim$$

Ahora, por el Axioma de Elección, podemos crear una aplicación $\phi : \mathbb{R}/\sim \to \mathbb{R}$ que de cada clase de equivalencia nos elija un elemento en $[0,1]$, por lo que componiendo $\phi \circ \pi$, obtenemos una aplicación:

$$\mathbb{R} \xrightarrow{\pi} \mathbb{R}/\sim \xrightarrow{\phi} \mathbb{R}$$

Y tomando $V = (\phi \circ \pi)(\mathbb{R})$, tenemos un conjunto de Vitali en $[0,1]$. □

3.4. Jerarquía generalizada

Con el fin de que cualquier subconjunto que consideremos esté en algún lugar de la jerarquía, tratamos de extener el concepto de los ordinales finitos (los números naturales) a otros ordinales más generosas. Para ello, recordamos cómo se construyeron los números naturales:

$$\begin{aligned} 0 &= \emptyset \\ 1 &= \{\emptyset\} \\ 2 &= \{\emptyset, \{\emptyset\}\} \\ 3 &= \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} \\ &\vdots \end{aligned}$$

Es decir, a partir del conjunto $\emptyset$, cuya existencia suponemos por axioma, y con la aplicación sucesor s, dada por:

$$s(x) = x \cup \{x\}$$

De esta forma, tenemos que:

$$\begin{aligned} 2 &= 0 \cup 1 \\ 3 &= 0 \cup 1 \cup 2 \\ &\vdots \end{aligned}$$

Y si queremos generalizar estos conceptos para que transciendan a los naturales, tras todos los números naturales, el siguiente ordinal a considerar será $\mathbb{N}$, luego:

$$\mathbb{N} + 1 := \{\mathbb{N}, \{\mathbb{N}\}\} \quad \dots$$

Con esta idea, llegamos a la siguiente definición:

Definición 3.5 (Ordinal)**.** Un conjunto α es un ordinal cuando:

1. Es transitivo para $\in$:

$$\forall x(x \in \alpha \implies \forall y(y \in x \implies y \in \alpha))$$

2. α es bien ordenado para $\in$.

3.4.1. Motivación

Resulta que la jerarquía que hemos construido no es suficiente, puesto que hay conjuntos Borelianos que se salen de la jerarquía.

Definición 3.6. Diremos que un punto de un espacio topológico $x \in X$ es aislado si $\{x\}$ es abierto.
Además, consideraremos:

$$S(X) = \{x \in X \mid x \text{ no es aislado}\}$$

Ejemplo. Veamos algunos ejemplos de esto:

- Si consideramos:
$$D = \left\{1 - \frac{1}{n} \mid n \in \mathbb{N}\right\} \cup \{1\}$$
Resulta que
$$S(D) = \{1\} \quad S(S(D)) = \{\emptyset\} \quad S^3(d) = S(S(S(D))) = \emptyset$$
Vemos que "D se estabiliza en 2 iteraciones".

- Si consideramos ahora:
$$E = \left\{1 - \frac{1}{n_1} \mid n_1 \in \mathbb{N}\right\} \cup \left\{2 - \frac{1}{n_1} - \frac{1}{n_2} \mid n_1, n_2 \in \mathbb{N}\right\} \cup \{2\}$$
Tendremos:
$$S(E) = \{1\} \cup \{2 - \frac{1}{n_2} \mid n_2 \in \mathbb{N}\} \quad S(S(E)) = \{2\} \quad S^3(E) = \emptyset$$

- Dado $m \in \mathbb{N}$:
$$X_m = \bigcup_{k=1}^{m} \left\{k - \frac{1}{n_1} - \ldots - \frac{1}{n_k} \mid n_1, \ldots, n_k \in \mathbb{N}\right\} \cup \{m\}$$
Tendremos:
$$S^m(X_m) = \{m\} \qquad S^{m+1}(X_m) = \emptyset$$
Por último, si consideramos:
$$X = \bigcup_{n \in \mathbb{N}} X_n$$
Llegaremos a que $S^m(X) \neq \emptyset \ \forall m \in \mathbb{N}$. Sabemos que $]0,1[$ es homeomorfo a $]0,+\infty[$. Pues bien, si aplicamos dicho homeomorfismo a X, obtendremos que $\{1\} \in S^m(X)$, $\forall m \in \mathbb{N}$.

Veamos unos últimos luego entenderemos mejor y que nos serirán de motivación:

Teorema 3.5. *Si* $\sum_{n=-\infty}^{\infty} c_n e^{int} = 0$ *para todo* $t \in [-\pi, \pi]$*, necesariamente ha de cumplirse que* $c_n = 0$ *para todo* $n \in \mathbb{Z}$.

Teorema 3.6 (Dedekind). *Si* $\sum_{n=-\infty}^{\infty} c_n e^{int} = 0$ *para todo* $t \in [-\pi, \pi]$ *salvo en un número finito de puntos, necesariamente ha de cumplirse que* $c_n = 0$ *para todo* $n \in \mathbb{Z}$.

Dedekind preguntó a Cantor cómo mejorar las hipótesis del Teorema, algo que consiguió y que luego entenderemos, ya que son necesarias algunas definiciones previas.

Teorema 3.7. *Si* $\sum_{n=-\infty}^{\infty} c_n e^{int} = 0$ *para todo* $t \in [-\pi, \pi]$ *salvo en un conjunto* X *"que se estabilice en una etapa numerable", necesariamente ha de cumplirse que* $c_n = 0$ *para todo* $n \in \mathbb{Z}$.

3.5. Buenos órdenes

Definición 3.7. Una relación[8] $<$ es un orden en un conjunto X si:

- $<$ es anti-reflexiva ($\forall x \in X\ \neg x < x$).
- $<$ es transitiva ($\forall x, y, z \in X\ x < y < z \to x < z$).

Diremos que:

- $<$ es un orden total si $\forall x, y \in X\ (x < y \vee x = y \vee y < x)$.
- $<$ es un buen orden si $\forall Y \subset X\ (Y \neq \emptyset \to Y$ tiene un mínimo).

Diremos que una aplicación $f : (X, <) \to (Y, <)$ entre dos conjuntos bien ordenados (es decir, que tienen un buen orden) es un morfismo si:

$$\forall x, x' \in X\ (x < x' \to f(x) < f(x'))$$

Diremos que:

- f es un isomorfismo si es un morfismo biyectivo con inversa un morfismo.
- f es un automorfismo si es un isomorfismo con mismo dominio que codominio.

Resulta que la condición de tener un buen orden es mucho más general que la de tener un orden total.

Proposición 3.8. *Todo conjunto bien ordenado tiene un orden total.*

Demostración. Sea W un conjunto bien ordenado, sean $x, y \in W$, entonces el conjunto $\{x, y\} \subseteq W$ tiene un mínimo, es decir:

- Bien $x \leqslant y$.
- Bien $y \leqslant x$.

Por lo que se cumple la definición del orden total para el orden de W. □

[8] Puede hacerse la definición para $<$ o para $\leqslant$.

Lema 3.9. *Si $(W, <)$ es un buen orden y $f : W \to W$ es un morfismo, entonces:*

$$f(x) \geqslant x \qquad \forall x \in W$$

Demostración. Como la mayoría de las demostraciones que involucren a un buen orden: lo probaremos por reducción al absurdo, cogiendo el mínimo elemento del conjunto en el que no se cumple la tesis.

Por reducción al absurdo, suponemos que:

$$X = \{x \in W \mid f(x) < x\}$$

es no vacío. Por ser W bien ordenado, este ha de tener un mínimo, $y = \min X$. Por ser $y \in X$, tendremos que $f(y) < y$. Si volvemos a aplicar f:

$$f(f(y)) < f(y) < y$$

Sea $z = f(y)$, tenemos que $z \in X$ por ser $f(z) < z$ y además tenemos que $z < y$, por lo que y no podía ser el mínimo de X, contradicción, que venía de suponer que X era no vacío. □

Ejemplo. Sin embargo, el lema anterior no se verifica si solo consideramos que $(W, <)$ tenga un orden total. Por ejemplo, $(\mathbb{Z}, <)$ con el orden al que estamos acostumbrados es un orden total que no es un buen orden y podemos considerar el morfismo:

$$\begin{array}{rccc} f: & \mathbb{Z} & \longrightarrow & \mathbb{Z} \\ & z & \longmapsto & z-1 \end{array}$$

Corolario 3.9.1. *El único automorfismo de un buen orden es la identidad.*

Demostración. Sea $(W, <)$ un conjunto bien ordenado, sea $f : W \to W$ un automorfismo, si $x \in X$, tendremos por el Lema anterior que $f(x) \geqslant x$. Sin embargo, por ser f un autormofismo, también lo será f^{-1}, por lo que también tendremos que:

$$x = f^{-1}(f(x)) \geqslant f(x)$$

De donde concluimos que $f(x) = x$, para todo $x \in X$. □

Corolario 3.9.2. *Si W_1, W_2 son dos buenos órdenes isomorfos, el isomorfismo es único.*

Demostración. Supongamos que $f_1 : W_1 \to W_2$ y $f_2 : W_1 \to W_2$ son dos isomorfismos. Entonces, tendremos que $(f_2^{-1} \circ f_1) : W_1 \to W_1$ es un automorfismo, por lo que debe ser:

$$f_2^{-1} \circ f_1 = id \Longrightarrow f_1 = f_2 \circ f_2^{-1} \circ f_1 = f_2$$

□

Definición 3.8. Sea W un buen orden, para $x \in W$ definimos el segmento inicial generado por x como:

$$W(x) = \{y \in W \mid y < x\}$$

Es claro que $W(x)$ también es un buen orden.

Lema 3.10. *Ningún buen orden es isomorfo a un segmento inicial propio suyo.*

Demostración. Sea W un buen orden, supongamos que tenemos $W \cong W(x)$ para cierto $x \in W$. En dicho caso, si $f : W \to W(x)$ es un isomorfismo, tendremos:

- $f(x) \geqslant x$ por ser f un morfismo.
- $f(x) < x$ por ser $f(x) \in W(x)$.

Hemos llegado a una contradicción. □

Teorema 3.11. *Sean W_1 y W_2 buenos órdenes, entonces ocurre uno y solo uno de los siguientes escenarios:*

- $W_1 \cong W_2$.
- $W_1 \cong W_2(x)$ *para cierto* $x \in W_2$.
- $W_2 \cong W_1(x)$ *para cierto* $x \in W_1$.

Demostración. Consideramos:

$$f = \{(x, y) \in W_1 \times W_2 \mid W_1(x) \cong W_2(y)\}$$

- En primer lugar, veamos que dado $x \in W_1$, existe un único $y \in W_2$ de forma que $(x, y) \in f$. Para ello, sean $x \in W_1$, $y_1, y_2 \in W_2$ de forma que $(x, y_1), (x, y_2) \in f$, entonces tendremos que $W_1(x) \cong W_2(y_1)$ y que $W_1(x) \cong W_2(y_2)$. Por la transitividad de la isomorfía, llegamos a que $W_2(y_1) \cong W_2(y_2)$ y como ningún buen orden puede ser isomorfo a un segmento inicial propio, concluimos que ha de ser $y_1 = y_2$.
- Ahora, veamos que dados $x_1, x_2 \in W_1$ de forma que si $y \in W_2$ con $(x_1, y), (x_2, y) \in f$, entonces $x_1 = x_2$.

  La demostración es análoga a la anterior.
- Es fácil ver que f es un morfismo.

Supongamos ahora que $dom(f) \neq W_1$ y que $ran(f) \neq W_2$. Consideramos por tanto, $x = min(W_1 \setminus dom(f))$, $y = min(W_2 \setminus ran(f))$, tenemos que:

$$dom(f) = W_1(x) \qquad ran(f) = W_2(f)$$

Por la definición de f, tendremos que $(x, y) \in f$, que es una contradicción.

Como la hipótesis de la que partíamos era falsa, a de ser falsa alguna de las premisosas de la conjunción, o bien ambas:

- Si ambas son falsas, estamos en la primera situación.
- Si la primera es falsa, $W_1 \cong W_2(x)$ para cierto $x \in W_2$.
- Si la segunda es falsa, $W_2 \cong W_1(x)$ para cierto $x \in W_1$.

Falta ver que si tenemos $f : W_1 \to Y \subset W_2$, entonces si tomamos $x = \text{mín}(W_2 \setminus Y)$, tendremos que $Y = W_2(x)$, por lo que Y es un segmento inicial. □

Definición 3.9 (Ordinal)**.** Un ordinal es una clase de equivalencia de ser isomorfos entre buenos órdenes.

De esta forma, como ordinales tenemos:

$$0,\ 1,\ 2,\ \ldots\ n,\ \ldots\ \mathbb{N},\ \mathbb{N}+1,\ \ldots\ \mathbb{N}+n,\ \ldots\ \mathbb{N}+\mathbb{N},\ldots$$

Definición 3.10 (Cardinal)**.** Un cardinal es un ordinal que no está en biyección con ningún otro ordinal menor.

Notación. En este ámbito nos encontramos con dos notaciones equivalentes:

- ω_0
- $\aleph_0$

Ambos representan a $\mathbb{N}$. Una pregunta que se hizo Cantor es cómo conseguir el siguiente cardinal, ω_1 o $\aleph_1$, y si este era el mismo que el de $\mathbb{R}$.

Paul Cohen demostró que ver si $\aleph_1$ coincide con el cardinal de $\mathbb{R}$ es algo que no se puede demostrar, ya que hay modelos a partir de los axiomas de Zermelo-Fraenkel que sí lo cumplen y otros modelos que no.

3.6. Universalidad

El concepto de universalidad es extendido en matemáticas y depende del área en el que nos encontremos. En general, diremos que un espacio (un conjunto con una estructura) es universal para una "clase de espacios" si:

- Dicho conjunto pertenece a la clase.
- Contiene "una copia" (depende del contexto en el que nos encontremos será una clase u otra) de todos los elementos de la clase.

3.6.1. Espacios topológicos metrizables y separables

El objetivo de esta sección es probar que:

Teorema 3.12. *El cubo de Hilbert $I = [0,1]^{\mathbb{N}}$ es universal (bajo homeomorfismos) en la clase de los espacios métrizables y separables.*

Para probar el Teorema, hemos de probar:

- Primero, que el cubo de Hilbert es un espacio metrizable y separable.
- Segundo, que todo espacio metrizable y separable contiene una copia homeomorfa dentro del cubo de Hilbert.

En primer lugar, sobre el cubo de Hilbert podemos definir la distancia:

$$d(x,y) = \sum_{n=0}^{+\infty} \frac{d(x_n, y_n)}{2^n} \qquad \forall x, y \in [0,1]^{\mathbb{N}}$$

Ejercicio. Demostrar que d es una métrica en $[0,1]^{\mathbb{N}}$.

Lema 3.13. *$I = [0,1]^{\mathbb{N}}$ es separable.*

Demostración. Lo primero que podríamos pensar es considerar el conjunto $(\mathbb{Q} \cap [0,1])^{\mathbb{N}} \subseteq I$ como nuestro conjunto denso y numerable. Sin embargo, este conjunto no es numerable, ya que:

$$2^{\mathbb{N}} \subseteq (\mathbb{Q} \cap [0,1])^{\mathbb{N}}$$

Lo que podemos hacer es tomar una cantidad finita de elementos de $\mathbb{Q} \cap [0,1]$ y seguir completando la sucesión con ceros.

Sea $X = \{x_n \mid n \in \mathbb{N}\}$ un conjunto denso de $[0,1]$ (por ejemplo, $\mathbb{Q} \cap [0,1]$, aunque valdría cualquier otro), definimos:

$$Y_n = \{y \in I \mid y_0, \ldots, y_n \in X \wedge y_m = 0 \quad \forall m \geqslant n\}$$

Y consideramos:

$$Y = \bigcup_{n \in \mathbb{N}} Y_n$$

Que es numerable por ser unión numerable de conjuntos numerables. Veamos que Y es denso en I. Sea $z \in I$, y $\varepsilon > 0$, buscamos $y \in Y$ de forma que $d(z,y) < \varepsilon$. Para ello, sea $n_0 \in \mathbb{N}$ de forma que:

$$\sum_{n=n_0}^{+\infty} \frac{1}{2^n} \leqslant \frac{\varepsilon}{2}$$

Lo que haremos será buscar (hágase) $y_0, \ldots, y_{n_0} \in X$ de forma que:

$$\sum_{n=1}^{n_0} \frac{d(y_n, z_n)}{2^n} \leqslant \frac{\varepsilon}{2}$$

De esta forma, tendremos que:

$$y = (y_0, y_1, \ldots, y_{n_0}, 0, \ldots, 0, \ldots) \in Y$$

Así como que:

$$d(z,y) \leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

□

Estamos ya en condiciones de demostrar el Teorema anterior:

Teorema 3.14. *El cubo de Hilbert $I = [0,1]^{\mathbb{N}}$ es universal (bajo homeomorfismos) en la clase de los espacios métrizables y separables.*

Demostración. Hemos probado ya que I es un espacio metrizable y separable. Sea ahora (X, d) un espacio métrico y separable, queremos definir una función $f : X \to [0,1]^{\mathbb{N}}$ continua, abierta e inyectiva, para que la restricción a su imagen sea un homeomorfismo.

En la asignatura de Análisis Matemático I se demostró que siempre que tengamos un espacio métrico (X, d), podremos definir la distancia:

$$d'(x,y) = \frac{d(x,y)}{d(x,y)+1} \qquad \forall x, y \in X$$

Que también será una distancia, equivalente a d y acotada por 1:

$$d'(x,y) \leqslant 1 \qquad \forall x, y \in X$$

De esta forma, por ser X un espacio separable, admitirá un conjunto denso y numerable, cuyos elementos denotaremos por x_n. Podemos ahora definir f por:

$$f(x) = \{d'(x, x_n)\}_{n\in\mathbb{N}} \qquad \forall x \in X$$

- Para probar la inyectividad de f, sean $x, y \in X$ distintos, tomamos $r = d'(x,y)$ y $x_n \in B(x, r/4) \cap X$, tenemos entonces que $d(x, x_n) \neq d(y, x_n)$, por lo que:

$$f(x) = \{d'(x, xn)\}_{n\in\mathbb{N}} \neq \{d'(y, x_n)\}_{n\in\mathbb{N}} = f(y)$$

- La continuidad de la función se deja como ejercicio.

- Para demostrar que f es abierta, lo que haremos será probar que f^{-1} es continua. Para ello, sea $\{f(y_n)\}_{n\in\mathbb{N}} \subseteq Im(f)$ de forma que $\{f(y_n)\} \to f(y)$, queremos probar que $\{y_n\} \to y$. Observemos que tenemos:

$$\{d(x_n, y_m)\}_{m\in\mathbb{N}} \overset{n}{\to} \{d(y, x_m)\}_{m\in\mathbb{N}}$$

  Es decir, para cada $m \in \mathbb{N}$ tenemos que $\{d(y_n, x_m)\} \to d(y, x_m)$. Sea $\varepsilon > 0$, buscamos $n_0 \in \mathbb{N}$ para el cual si $n \geqslant n_0$, entonces $d(y, y_n) \leqslant \varepsilon$:

$$d(y, y_n) \leqslant d(y, a) + d(y_n, a) \leqslant \varepsilon$$

□

3.6.2. Conjuntos universales para la jerarquía

De vuelta a la Jerarquía Boreliana, si consideramos el conjunto de Cantor $\mathcal{C} = 2^{\mathbb{N}}$, definimos lo que es un conjunto universal para una clase de dicha jerarquía:

Definición 3.11. Fijado un conjunto X, decimos que $\mathcal{U}$ es $\mathcal{C}-$universal para $\Sigma^0_\alpha(X)$ para cierto α si:

- $\mathcal{U} \in \Sigma^0_\alpha(\mathcal{C} \times X)$.
- Para todo $Y \subseteq X$, $Y \in \Sigma^0_\alpha \iff \exists c \in \mathcal{C}$ con $Y = \pi_{\mathcal{C}}(\mathcal{U})$.

Lema 3.15. *Sea (X, d) un espacio polaco, existe $\mathcal{U}$, un conjunto $\mathcal{C}-$universal para los abiertos de X.*

Demostración. Al ser X separable, existe $\{O_n\}_{n\in\mathbb{N}}$ una familia de abiertos básicos de forma que para todo O abierto, existe $\sigma : \mathbb{N} \to \mathbb{N}$ de forma que:

$$O = \bigcup_{n\in\mathbb{N}} O_{\sigma(n)}$$

Es decir:

$$O \in \Sigma_1^0(X) \iff \exists c \in \mathcal{C} \text{ con } O = \bigcup_{n\in\mathbb{N}} O_n$$

De esta forma, sea:

$$\mathcal{U} = \{(c, x) \in 2^{\mathbb{N}} \times X \mid \exists n \in \mathbb{N}(c(n) = 1 \wedge x \in On)\}$$

Tenemos que:

$$Y \in \Sigma_1^0(X) \iff \exists c \in \mathcal{C} \text{ con } Y = \pi_{\mathcal{C}}(\mathcal{U})$$

□

Lema 3.16. *Si $\mathcal{U}_\alpha$ es $\mathcal{C}-$universal para $\Sigma_\alpha^0(X)$, entonces existe $\mathcal{U}_{\alpha+1}$, un conjunto $\mathcal{C}-$universal para $\Sigma^0 0_{\alpha+1}(X)$.*

Lema 3.17.

$$\Sigma_\alpha^0 \neq \Pi_\alpha^0$$

Demostración. Sea $\mathcal{U}_\alpha$ un conjunto $\mathcal{C}-$uniersal en $\Sigma_\alpha^0(\mathcal{C})$, supongamos que $\Sigma_\alpha^0 = \Pi_\alpha^0$, definimos el conjunto A como:

$$y \in A \iff (y, y) \notin \mathcal{U}_\alpha$$

De esta forma, tenemos que $\mathcal{U}_\alpha \in \Sigma_\alpha^0 = \Pi_\alpha^0$, por lo que $\exists c \in \mathcal{C}$ de forma que:

$$A = \pi_{\mathcal{C}}(\mathcal{U}_\alpha)$$

Si nos preguntamos si (c, c) está o no en A:

- Si $c \in A$, entonces $(c, c) \in \mathcal{U}_\alpha$, contradicción.
- Si $c \notin A$, entonces $(c, c) \notin \mathcal{U}_\alpha$, por lo que $c \in A$, contradicción.

3.7. Conjuntos completos

Este apartado nos servirá para comprobar si una cierta propiedad es de una complejidad concreta y no de ninguna anterior. La idea que usaremos es parecida a la que se usa en geometría o topología, con medidas invariantes por las transformaciones que nos interesan. Por ejemplo, si X, Y son espacios polacos y $A \subseteq X$, si $A \in \Sigma_2^0(X)$, entonces existe una cantidad numerable de cerrados A_n de forma que:

$$A = \bigcup_{n\in\mathbb{N}} A_n$$

Sea ahora $f : Y \to X$ una función continua, observemos que:

$$f^{-1}(A) = f^{-1}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \bigcup_{n\in\mathbb{N}} f^{-1}(A_n) \in \Sigma^0_2(Y)$$

Por inducción puede probarse para todo Σ^0_n, así como para las clases Π^0_n. Esta propiedad fomenta la siguiente definición:

Definición 3.12. Sean X, Y espaciones polacos, si $A \subseteq X$ y $B \subseteq Y$, diremos que A es wadge reducible (o simplemente reducible) a B ($A \leqslant_w B$) si existe una función $f : X \to Y$ continua de forma que:

$$f^{-1}(B) = A$$

Definición 3.13. Si Γ es una clase de la Jerarquía Boreliana (bien Σ^0_n bien Π^0_n para cierto n), un conjunto $B \subseteq Y$ de un espacio polace se dice que es $\Gamma-$copleto si:

- $B \in \Gamma(Y)$.
- Para todo espacio polaco X, si $A \in \Gamma(Y)$, entonces $A \leqslant_w B$.

De esta forma, si quiero saber la copmlejidad de un conjunto O del que solo conozco que $O \in \Gamma$, tendremos próximamente una variedad de ejemplos sencillos que sabremos que son $\Gamma-$completos. Sea U un conjunto $\Gamma-$completo, tendremos entonces que existirá una función continua f con:

$$f^{-1}(U) = O \qquad O \leqslant_w U$$

Y nos bastará demostrar que $U \leqslant_w O$ para ver que el orden de complejidad de O es exactamente Γ.

3.7.1. Problema de Cauchy

Nos interesaremos por un ejemplo concreto, el de clasificar la complejida de las EDOs (Ecuaciones Diferenciales Ordinarias) con valores iniciales que tienen soluciones únicas.

Para dar una EDO, necesitamos dar:

- Una función $F : \mathbb{R}^2 \to \mathbb{R}$ continua.
- Una condición inicial $(t_0, x_0) \in \mathbb{R}^2$.

Y lo que exigiremos a una función $\varphi : \mathbb{R} \to \mathbb{R}$ continua, será:

$$\begin{cases} x' = F(t, x) \\ x(t_0) = x_0 \end{cases}$$