

MATERIA
Teoría de Autómatas e Linguaxes Formais

unidade
didáctica
1

TITULACIÓN
Grao en Enxeñaría Informática

Autómatas de estados finitos

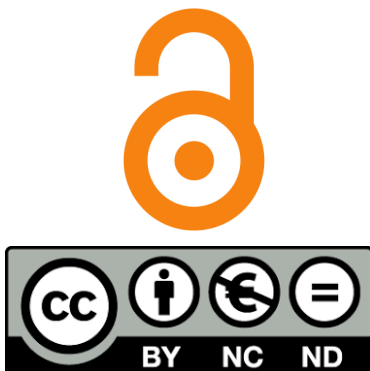
Daniel Cores Costa

Área de Ciencias da Computación e Intelixencia Artificial
Electrónica e Computación
Escola Técnica Superior de Enxeñaría

unidadesdidácticas
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA



Ensinanzas Técnicas



Esta obra atópase baixo unha licenza internacional Creative Commons BY-NC-ND 4.0. Calquera forma de reprodución, distribución, comunicación pública ou transformación desta obra non incluída na licenza Creative Commons BY-NC-ND 4.0 só pode ser realizada coa autorización expresa dos titulares, salvo excepción prevista pola lei. Pode acceder Vde. ao texto completo da licenza nesta ligazón: <https://creativecommons.org/licenses/by-nc-nd/4.0/deed.gl>

© Universidade de Santiago de Compostela, 2021

Deseño e maquetación

J. M. Gairí

Edita

Edicións USC

usc.es/publicacions

DOI

<https://dx.doi.org/10.15304/9788419155153>



MATERIA: Teoría de Autómatas e Linguaxes Formais

TITULACIÓN: Grao en Enxeñaría Informática

PROGRAMA XERAL DO CURSO

Localización da presente unidade didáctica

Unidade Didáctica I: Autómatas de estados finitos

Autómatas de estados finitos deterministas

Autómatas de estados finitos non deterministas

Clausuras respecto da cadea baleira

Equivalencia entre autómatas de estados finitos deterministas e non deterministas

Equivalencia de linguaxes regulares

Unidade Didáctica II: Linguaxes regulares

Unidade Didáctica III: Gramáticas independentes de contexto

Unidade Didáctica IV: Autómatas con pila

Unidade Didáctica V: Máquinas Turing

Unidade Didáctica VI: Decidibilidade e complexidade



ÍNDICE

CONTEXTUALIZACIÓN

1. Presentación
2. Xustificación
3. Obxectivos e Competencias
4. Metodoloxía
 - 4.1. Sesións expositivas
 - 4.2. Sesións interactivas

CONTIDOS

1. Conceptos previos
2. Autómatas de estados finitos deterministas
 - 2.1. Definición
 - 2.2. Diagrama de transicións
 - 2.3 Táboa de transición
 - 2.4. Función de transición estendida
3. Autómatas de estados finitos non deterministas
 - 3.1. Definición
 - 3.2. Función de transición estendida
 - 3.3. Autómatas con cadea baleira
4. Clausuras respecto da cadea baleira
5. Equivalencia entre autómatas de estados finitos deterministas e non deterministas
 - 5.1. Construción de subconxuntos
6. Minimización dun autómata de estados finitos deterministas
7. Equivalencia de linguaxes regulares

AVALIACIÓN

BIBLIOGRAFÍA



CONTEXTUALIZACIÓN

1. Presentación

A unidade didáctica “Autómatas de estados finitos” enmárcase na materia Teoría de Autómatas e Linguaxes Formais que se imparte no primeiro cuadrimestre do terceiro curso do Grao en Enxeñaría Informática da Universidade de Santiago de Compostela. Esta materia ten carácter obrigatorio e conta cunha carga lectiva de 6 créditos ECTS.

Ademais da exposición dos contidos teóricos, nesta UD tamén se contemplan dúas sesións prácticas de 2 horas. A primeira está centrada na resolución de problemas de forma autónoma por parte do alumnado. Na segunda, abordaranse as aplicacións prácticas dos contidos expostos.

2. Xustificación

Nesta unidade didáctica abórdase o primeiro bloque da materia centrado en autómatas de estados finitos. Trátase dos dispositivos de computación teóricos máis sinxelos. A súa introdución servirá para asentar os conceptos básicos que permitirán avanzar ao longo do curso presentando modelos máis complexos de forma iterativa.

Esta materia supón a primeira toma de contacto co módulo de Sistemas Intelixentes. Polo tanto, esta unidade didáctica non serve só como introdución á Teoría de Autómatas e Linguaxes Formais, senón que se trata dunha introdución a este bloque. Polo tanto, os coñecementos e destrezas adquiridos serán de grande importancia para abordar con éxito as demais materias desta área de coñecemento. Ademais, os contidos da materia, e en particular os recollidos nesta unidade didáctica, serán de vital importancia para materias alleas a este bloque como a materia de Compiladores e Intérpretes impartida no cuarto curso. De forma xeral, trátanse conceptos fundamentais para o procesamento de textos. Polo que os conceptos introducidos aparecen de forma recorrente tanto en materias relacionadas co deseño de interfaces de usuario como coa análise do funcionamento de sistemas operativos.

3. Obxectivos e Competencias

As seguintes competencias globais recollidas na memoria da titulación abórdanse nesta materia:

- **CG9:** capacidade de resolución de problemas con iniciativa, creatividade e toma de decisión de forma autónoma. Capacidade de comunicación dos coñecementos, habilidades e destrezas adquiridos.
- **TR1:** capacidade de análise e síntese. Capacidade de organización e planificación. Comunicación oral e escrita en galego, castelán e inglés. Capacidade de xestión da información. Resolución de problemas. Toma de decisións.
- **TR3:** aprendizaxe de forma autónoma. Adaptación a novas circunstancias. Creatividade. Iniciativa e espírito crítico emprendedor. Motivación pola calidade. Sensibilidade cara a temas medioambientais.



A continuación recóllense as competencias específicas asociadas co título de Grao en Enxeñaría Informática e Enxeñaría Técnica Informática establecidas polo Consello de Universidades (acordo 03/03/2009):

- **FB3:** capacidade para comprender e dominar os conceptos básicos de matemática discreta, lóxica, algorítmica e complexidade computacional, e a súa aplicación para a resolución de problemas propios da enxeñaría.
- **RI6:** coñecemento e aplicación dos procedementos algorítmicos básicos das tecnoloxías informáticas para o deseño de solucións a problemas, analizando a idoneidade e complexidade dos algoritmos propostos.

A materia está deseñada para acadar os seguintes obxectivos xerais:

- Coñecer e comprender as posibilidades dos diferentes modelos abstractos de computación.
- Coñecer e comprender os fundamentos e os límites da computación, así como a distinción entre problemas computables e non computables, distinguindo entre problemas tratables e intratables dentro dos primeiros.
- Coñecer, comprender e saber aplicar un conxunto de técnicas da investigación de operacións que dan lugar a algoritmos coñecidos dentro da resolución de problemas de enxeñaría.
- Coñecer, comprender e manexar técnicas de resolución de problemas baseados en representación do coñecemento.
- Coñecer, comprender e saber aplicar metodoloxías de deseño de sistemas baseados en coñecemento.

Os seguintes obxectivos específicos serán abordados a través dos contidos desta UD:

- Coñecer e comprender os modelos de computación máis básicos: os autómatas de estados finitos.
- Coñecer e comprender as posibles aplicacións dos autómatas de estados finitos na resolución de problemas de enxeñaría.
- Coñecer, comprender e saber aplicar técnicas de deseño de autómatas de estados finitos para a representación de coñecemento.

4. Metodoloxía

A metodoloxía docente organízase en sesións expositivas e sesións interactivas de resolución de casos prácticos.

4.1. Sesións expositivas

A estas sesións asiste o total do alumnado matriculado na materia. A organización destas sesións conta con dúas partes diferenciadas:

- Exposición dos contidos teóricos por parte do docente. Ao comezo de cada nova sección, o docente encargarase de introducir os principios teóricos da mesma. Para tal efecto, utilizaranse proxeccións con esquemas e diagramas de funcionamento dos diferentes modelos introducidos así como casos reais de aplicación. Deste xeito, búscase motivar ao alumnado a afondar nos contidos da materia. Para isto, tamén se indicarán os

capítulos específicos da bibliografía recomendada nos que se poden ampliar os conceptos introducidos nas sesións expositivas.

- Resolución de exercicios involucrando ao alumnado de forma activa. Co fin de asentar os coñecementos adquiridos, propóñense unha serie de exercicios teórico-prácticos que se resolven de forma grupal nas propias sesións. Búscase que o alumnado propoña ideas acerca de como abordar os problemas presentados polo docente. A asimilación por parte do alumnado dos conceptos expostos será de vital importancia para o correcto desenvolvemento das sesións interactivas.

4.2. Sesións interactivas

A presente unidade didáctica conta con dúas sesións interactivas en grupos reducidos dentro do programa da materia. Nestas sesións búscase unha atención máis personalizada ao alumnado facilitada polo reducido número de alumnos por grupo. Os temas que se van tratar en cada unha das sesións son os seguintes:

- Resolución de problemas de forma autónoma. Preséntanse problemas relacionados cos contidos expostos nos seminarios. Os problemas divídense entre o alumnado presente na aula organizado en grupos. Cada grupo debe propoñer unha idea inicial acerca de como abordar cada un dos problemas e presentarllo os seus compañeiros na aula. Foméntase desta forma a presentación de contidos por parte do alumnado. Tamén se fomenta que os propios compañeiros propoñan alternativas propiciando un pequeno debate acerca das solucións propostas. A continuación da posta en común de ideas, búscase unha resolución completa individual de todos os exercicios que poida favorecer a aparición de novas dúbidas que requiran unha atención máis personalizada.
- Aplicacións prácticas. Propóñense unha serie de problemas reais para os cales a utilización de autómatas de estados finitos pode ser unha solución axeitada. Búscase mostrar a aplicación práctica dos contidos expostos e traballados nas sesións anteriores. A implementación das solucións que resolvan os problemas presentados polo docente debe facerse de forma individual. Os exercicios están deseñados de tal forma que os alumnos teñen que dedicar tempo de forma autónoma fora de clase. De todos xeitos, ofrécese a posibilidade de concertar titorías para abordar os posibles problemas que poidan xurdir durante este traballo autónomo.

CONTIDOS

1. Conceptos previos

Existen tres conceptos básicos da teoría de linguaxes formais imprescindibles para a comprensión da unidade didáctica, e da materia en xeral. Trátase dos alfabetos, as palabras e as linguaxes:



- Os alfabetos están formados por un conxunto non baleiro de símbolos.
- As palabras defínense como unha secuencia finita de símbolos pertencentes a un alfabeto. A palabra ou cadea baleira é un caso especial que consta de cero símbolos e denótase habitualmente por λ ou ϵ .
- As linguaxes están formadas por un número finito ou infinito de cadeas definidas sobre un alfabeto.

A Táboa 1 recolle as operacións definidas sobre palabras.

Táboa 1: Operacións con palabras

Operación	Representación	Definición
Concatenación	xy	Dadas dúas palabras x e y , o resultado será unha palabra con todos os símbolos de x seguidos polos símbolos de y .
Potencia	x^i	Dada unha palabra x , a potencia i -ésima calcúlase como a concatenación de x i veces.
Reflexión	x^{-1}	Dada unha palabra x con símbolos $A_1 \dots A_n$, a súa reflexión calcúlase invertendo a orde dos símbolos que a forman: $A_n \dots A_1$.

A Táboa 2 recolle as operacións que se poden aplicar sobre linguaxes.

Táboa 2: Operacións con linguaxes

Operación	Representación	Definición
Unión	$L_1 \cup L_2$	A linguaxe resultado contén todas as palabras de L_1 e L_2 .
Intersección	$L_1 \cap L_2$	A linguaxe resultado contén todas as palabras que pertencen a L_1 e L_2 .
Resta	$L_1 - L_2$	A linguaxe resultado contén todas as palabras que pertencen a L_1 pero non a L_2 .
Concatenación	$L_1 \cdot L_2$	A linguaxe resultado contén todas as palabras que se poden formar concatenando palabras de L_1 con palabras de L_2 .
Potencia	L_1^i	A linguaxe resultado contén todas as palabras resultado de aplicar a potencia i -ésima as palabras de L_1 .
Clausura	L_1^*	A linguaxe resultado consiste na unión de todas as potencias de L_1 .
Clausura positiva	L_1^+	A linguaxe resultado consiste na clausura de L_1 engadindo a palabra baleira.
Reflexión	L_1^{-}	A linguaxe resultado está formada pola reflexión das palabras de L_1 .

Os autómatas son dispositivos teóricos capaces de recibir e transmitir información. Dada unha cadea de símbolos de entrada, producen una saída, en función de dita entrada e dos estados internos polos que transita o autómata.

En concreto, nesta unidade didáctica centrarémonos no funcionamento dos autómatas de estados finitos. Entre as principais aplicacións deste tipo de autómatas destacan:

- Implementacións software para o deseño e verificación do comportamento de circuitos dixitais.
- Analizador léxico dun compilador.
- Busca de patróns en textos.
- Comprobación do funcionamento de sistemas cun número finito de estados diferentes (protocolos de comunicación, protocolos de intercambio seguro de información...)

2. Autómatas de estados finitos deterministas

Un autómata de estados finitos está definido pola tupla (Q, S, δ, q_0, F) :

- Q : conxunto de estados.
- S : conxunto de símbolos de entrada.
- δ : función de transición. Existen dous tipos fundamentais de autómatas de estados finitos atendendo a función de transición: deterministas e non deterministas.
- q_0 : estado inicial.
- F : conxunto de estados finais.

Unha linguaxe L é recoñecida por un autómata se para todas as palabras da linguaxe se acadan un estado final do autómata. Os autómatas de estados finitos recoñecen as linguaxes regulares.

2.1. Definición

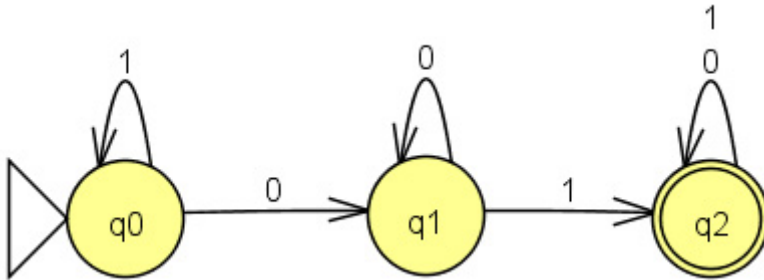
Os autómatas de estados finitos deterministas só poden transitar a un estado, dado un símbolo de entrada e partindo do estado actual. Polo tanto, a función de transición ten a forma $\delta(q, a)=p$, dado un estado q e un símbolo a transítase a un estado p .

2.2. Diagrama de transicións

De forma gráfica, un autómata pode representarse como un grafo tal e como se mostra na Figura 1. Os nodos do grafo representan o conxunto de estados Q . O estado inicial márcase cunha frecha (Estado q_0 na Figura 1). Os estados finais F represéntanse cun dobre círculo (Estado q_2 na Figura 1). O arco dun estado q a un estado p está definido por $\delta(q, a)=p$ etiquetándose co símbolo de entrada a .



Figura 1. Diagrama de transición dun autómata de estados finitos determinista



2.3 Táboa de transición

Como alternativa á representación gráfica mediante grafos, pódese empregar unha táboa de transicións. As filas representan os estados e as columnas os símbolos de entrada. A Táboa 3 define o autómata mostrado na Figura 1. Neste caso, os estados finais márcanse cun *.

Táboa 3. Táboa de transicións de un autómata de estados finitos determinista

	0	1
->q ₀	q ₁	q ₀
q ₁	q ₁	q ₂
*q ₂	q ₂	q ₂

2.4. Función de transición estendida

A función de transición estendida denótase por $\hat{\delta}$. Esta función de transición define a que estado se transita dende o estado actual dada unha cadea de símbolos en lugar dun único símbolo. Polo tanto, un autómata recoñece unha linguaxe L se para cada palabra w se alcanza un estado $\hat{\delta}(q_0, w)=p$ que pertence ao conxunto de estados finais F .

3. Autómatas de estados finitos non deterministas

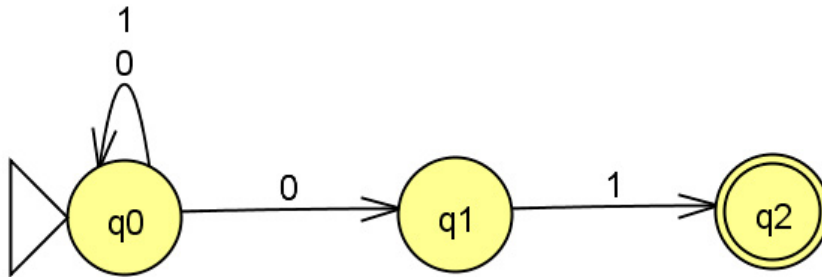
3.1. Definición

Ao contrario dos autómatas de estados finitos deterministas, no caso dos non deterministas, dado un símbolo de entrada pódese transitar a un conxunto



de estados. Este tipo de autómatas non engade ningunha linguaxe con respecto os recoñecidos polos autómatas de estados finitos deterministas. O non determinismo tan só permite a definición de autómatas máis compactos e facilmente interpretables. A Figura 2 mostra un autómata de estados finitos non determinista dado que, dende o estado q_0 lendo un 0 transítase aos estados q_1 e q_0 de forma simultánea.

Figura 2. Diagrama de transicións dun autómata de estados finitos non determinista



A Táboa 4 recolle a función de transición do autómata representado na Figura 2. Pódese ver como, a diferenza da Táboa 3, as celas da táboa conteñen conxuntos de estados en lugar de estados únicos. Cabe destacar que estes conxuntos de estados poden ser o conxunto baleiro \emptyset .

Táboa 4. Táboa de transicións de un autómata de estados finitos non determinista

	0	1
$\rightarrow q_0$	$\{q_0, q_1\}$	$\{q_0\}$
q_1	\emptyset	$\{q_2\}$
$*q_2$	\emptyset	\emptyset

3.2. Función de transición estendida

Ao igual que ocorre no caso dos autómatas de estados non deterministas, neste caso tamén se define unha función de transición estendida denotada por δ . De forma análoga, as linguaxes recoñecidas polo autómata serán aquelas para as que o conxunto de estados resultado de aplicar a todas as palabras da linguaxe contén como mínimo un estado final.



3.3. Autómatas con cadea baleira

Os autómatas de estados finitos non deterministas poden conter transicións coa cadea baleira. Este tipo de transicións non modifica o tipo de linguaxes que pode recoñecer un autómata de estados finitos. Simplemente, proporciona unha maior flexibilidade á hora de deseñar autómatas complexos dunha forma máis simple. A cadea baleira lese sempre de forma implícita, de tal xeito que cando se transita a un estado transítase de forma simultánea a todos os estados accesibles coa cadea baleira dende ese estado.

4. Clausuras respecto da cadea baleira

Un concepto fundamental cando se traballa coa cadea baleira é o de clausura dun estado q con respecto á cadea baleira: $CLAUS_{\varepsilon}(q)$. A $CLAUS_{\varepsilon}(q)$ contén todos os estados accesibles dende q empregando a cadea baleira. Para o cálculo de $CLAUS_{\varepsilon}(q)$ pode empregarse o seguinte algoritmo:

- **Paso base:** todos os estados accesibles dende q con ε pertencen a $CLAUS_{\varepsilon}(q)$.
- **Paso indutivo:** se un estado p está en $CLAUS_{\varepsilon}(q)$ engadir todos os estados pertencentes a $CLAUS_{\varepsilon}(p)$ á $CLAUS_{\varepsilon}(q)$.

Cando se traballa con autómatas de estados finitos non deterministas con cadea baleira, as clausuras deben terse en conta á hora de executar a función de transición. Dado un símbolo de entrada e un conxunto de estados actuais, o autómata transita a un novo conxunto de estados P . Realmente, dado que o autómata contén transicións coa cadea baleira, transita á unión das clausuras dos estados contidos en P .

5. Equivalencia entre autómatas de estados finitos deterministas e non deterministas

Toda linguaxe recoñecida por un autómata de estados finitos non determinista tamén pode ser recoñecida por un autómata de estados finitos determinista. O autómata determinista equivalente terá, como máximo, 2^n estados sendo n o número de estados do autómata de estados finitos non determinista.

O autómata de estados finitos determinista equivalente está definido por (Q, S, δ, q_0, F) . Onde o alfabeto de entrada S é o mesmo que no non determinista. O conxunto de estados Q será o conxunto de subconxuntos dos estados do autómata non determinista, sendo q_0 o conxunto de estados que conteña o estado inicial do autómata non determinista e F todos os conxuntos que conteñen estados finais no autómata de partida.

5.1. Construción de subconxuntos

A creación de conxuntos de estados do autómata de estados finitos non determinista utilízase para calcular os estados do autómata determinista equivalente. Cada estado do autómata resultado consistirá nun conxunto de estados do autómata



non determinista orixinal representando a posibilidade de estar en varios estados o mesmo tempo dunha forma determinista. Estes conxuntos poden calcularse seguindo un algoritmo de avaliación preguizosa:

- **Paso base:** créase un conxunto que contén a clausura do estado inicial do autómata de partida. Este conxunto será o estado inicial do autómata determinista.
- **Paso indutivo:** para cada conxunto de estados S calcúlase o conxunto de estados aos que se transita con cada símbolo do alfabeto $\delta(S, a)$. Os novos conxuntos representan os estados do autómata determinista. O proceso remata cando non se engaden novos conxuntos e xa se exploraron todos os símbolos do alfabeto para os conxuntos xa creados. Para a construción dos conxuntos débense ter en conta as clausuras xa que unha vez un autómata acadou un estado transítase de forma simultánea a todos os estados que se alcanzan mediante a cadea baleira.

6. Minimización dun autómata de estados finitos deterministas

Dado un autómata de estados finitos determinista, pódese atopar un autómata equivalente, que recoñece a mesma linguaxe, cun número de estados igual ou menor. Ademais, o autómata mínimo é único. Polo tanto, para comprobar se dous autómatas son equivalentes basta con calcular o autómata mínimo equivalente a cada un deles. Se estes autómatas mínimos son iguais, os autómatas de partida recoñecen a mesma linguaxe e polo tanto son equivalentes.

No proceso de minimización dun autómata de estados finitos determinista deben buscarse conxuntos de estados equivalentes. Deste xeito, o proceso de minimización terá como resultado unha serie de conxuntos de estados contendo todos os estados do autómata de partida. Os estados que pertencen ao mesmo conxunto son equivalentes entre si. Ademais, estados pertencentes a diferentes conxuntos non deben ser equivalentes, é dicir, deben ser distinguibles. A partición dos estados orixinais en clases de equivalencia denomínase conxunto cociente Q/E .

Cada un dos bloques de estados equivalentes representa un estado do autómata mínimo equivalente. Dado que, para cada símbolo do alfabeto, para todos os estados do mesmo conxunto se transita a estados pertencentes ao mesmo conxunto, pódese definir a nova función de transición analizando estas transicións.

7. Equivalencia de linguaxes regulares

Dadas dúas linguaxes L e M , para demostrar que son equivalentes é suficiente con comprobar a equivalencia entre os estados iniciais dos autómatas correspondentes que recoñecen as linguaxes L e M . Se os estados son equivalentes, dada calquera palabra de entrada, transítase a un estado de aceptación ou non aceptación nos dous casos. Polo tanto, os dous autómatas recoñecen a mesma linguaxe o que quere dicir que as linguaxes de partida son equivalentes.



AVALIACIÓN

A avaliación dos contidos expostos na UD farase tanto por medio de avaliación continua como incluíndo exercicios relacionados no exame final. A parte correspondente a avaliación continua está determinada polo resultado obtido nun exame realizado o rematar a UD. O contido deste exame tipo test está relacionado cos problemas resoltos nas sesións interactivas. Trátase dun test de resposta múltiple no que se penalizan as respostas erróneas.

O total da avaliación continua tendo en conta todas as actividades da manería contribúe a un 70 % da nota final mentres que o exame final ten un peso do 30 %. Será necesario acadar unha puntuación mínima de 5 nas dúas partes por separado para superar a materia. Os contidos da UD teñen un peso aproximado dun 10 % tanto da avaliación continua como do exame final.

BIBLIOGRAFÍA

J.E. HOPCROFT, R. MOTWANI e J.D. ULLMAN (2008), "Teoría de Autómatas, Lenguajes y Computación", Addison Wesley.

P. LINZ (2001), "An Introduction to Formal Languages and Automata", Jones and Bartlett Publishers.





Unha colección orientada a editar materiais docentes de calidade e pensada para apoiar o traballo do profesorado e do alumnado de todas as materias e titulacións da universidade

unidadesdidácticas
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

