

Arribat ST 22.10.90

29

UNIVERSITAT POLITÈCNICA DE CATALUNYA
DEPARTAMENT DE LLENGUATGES I SISTEMES INFORMÀTICS

TESI DOCTORAL

UN MODEL PEL TRACTAMENT DE LA INFORMACIÓ TEMPORAL
EN UN SISTEMA DE COMPRESIÓ AUTOMÀTICA DE NOTÍCIES

Memòria presentada per Núria Castell i Ariño
per optar al grau de Doctor en Informàtica
per la Universitat Politècnica de Catalunya,
sota la direcció de M^a Felisa Verdejo Maillo.

Barcelona, novembre de 1989

Als meus pares, que em varen obrir
les portes de la Universitat en moments
difícils i que sempre m'han fet costat.

Agraïments

El meu agraïment a:

- M^a Felisa Verdejo, sense la qual probablement no hauria seguit el camí de la recerca en intel·ligència artificial a la Universitat. Gràcies per la seva paciència i dedicació.
- M^a Antònia Martí, el seu ajut en el terreny lingüístic ha estat de gran valor i espero que la línia de recerca empresa continuï donant fruits.
- la Marta i en German, per la seva col·laboració desinterassada en tasques penoses i poc gratificants.
- Horacio Rodríguez, per compartir hores de programació i pel consol mutu davant les malifetes de l'Explorer.
- Josep Estrada, motivador del meu interès pel treball universitari, per la intel·ligència artificial i pel català. Ell va fer de pont amb un grup de recerca a l'empresa, i del treball amb aquest equip va nèixer la llavor de la meva tesi. Li dedico la redacció en català d'aquesta tesi.
- José Luis Balcázar, el Pepe, pel seu recolzament constant compartint nits, festius i vacances de feina i per la seva immensa contribució en el formatat d'aquesta memòria. La seva companyia ha convertit els llargs intervals de feina, cansament i malhumor en punts insignificants i ha transformat els punts de felicitat, els petits i escassos moments de tranquil·litat i trendresa, en valiosos intervals.
- Celestí Rosselló, pel seu ajut en el formatat final i sobretot pel seu recolzament moral.
- la Silvia i la Maite, pel seu oferiment reiterat de col·laboració i el seu ajut en detalls de la redacció.
- Kim Gabarró, per l'esforç de cuidar dels meus alumnes durant la darrera setmana.
- l'Esther, el Toni i el Vicenç, que han afegit al recolzament moral la seva preocupació per la salut de la doctoranda.
- tots els companys i amics que sempre m'han donat ànims. Em sap greu haver decebut als qui han ofert la seva col·laboració en vistes a la inexistent "traca final".
- al Departament pel seu esforç, malgrat les dificultats, en reduir la càrrega docent dels doctorands a base de passar-la als ja doctors. Gràcies a aquests companys per acceptar-ho.
- Caja Madrid, per la concessió d'una beca de finalització de tesi doctoral.
- Reko, Tima, Istmo, les marmotes, els pingüins, les foques, les sargantanes, els ocellets, els esquirolets, *los payasitos*, *los buhitos*, *las lapitas*, per la seva col·laboració simbòlica.

Índex

1. Introducció	1
2. Marc bibliogràfic	5
2.1 Sistemes de comprensió de textos sense tractament d'informació temporal	6
2.2 Modelització i representació de la informació temporal	9
2.3 Sistemes especialitzats en el tractament de la informació temporal	12
2.4 Sistemes de comprensió de textos amb tractament d'informació temporal	14
2.5 El sistema COTEM	17
3. Caracterització de l'aplicació: domini i exemple	19
3.1 Tipologia de les notícies	19
3.2 Tipologia dels verbs	25
3.3 Reflexions sobre la caracterització realitzada	27
3.4 Presentació de COTEM: un exemple	28
4. Informació temporal	33
4.1 Tipus d'informació temporal en el domini	33
4.2 Expressió lingüística del temps	34
4.2.1 Introducció	34
4.2.2 Assignació de l'estructura ternària (E-R-S) als temps verbals	36
4.2.3 Les expressions temporals: Tipologies	40
4.2.4 Relació entre expressions temporals, temps verbals i tipus de verbs	55
4.2.5 Conclusions	58
5. Components estructurals de COTEM	61
6. El Sistema Temporal	71
6.1 Components del Sistema Temporal	71
6.1.1 La Xarxa	72
6.1.2 El Calendari	76
6.2 Creació, propagació i consulta de relacions temporals	82

7. Funcionament de COTEM	89
7.1 El processador	89
7.1.1 Funcionament del processador	93
7.1.2 Processament de la informació temporal	98
7.2 L'incorporador	101
7.2.1 Tractament de la informació temporal per part de l'incorporador	104
7.3 Reflexions sobre el grau de generalitat	105
7.4 Exemple de funcionament	105
7.4.1 Una ampliació de capital	106
7.4.2 Uns resums econòmics	124
8. Conclusions	143
Annex: Verbs del domini i tipus associat	147
Bibliografia	161

1. Introducció

La comprensió del llenguatge natural ha passat per diferents períodes d'eufòria i desengany. Els primers objectius eren massa ambiciosos i els recursos lingüístics i informàtics bastant escassos, situació que va abocar al fracàs. Els grups d'investigació actuals s'enfronten amb plantejaments més realistes sobre la comprensió del llenguatge natural en l'àmbit, entre d'altres, de la construcció de bases de coneixement, un dels components fonamentals del sistema experts, la traducció automàtica i la creació d'interfícies en llenguatge natural per tal de facilitar la comunicació usuari-ordinador. Una àmplia exposició de tota la problemàtica es troba a [Allen-87].

Si bé en qüestions puntuals s'han produït avenços importants, queda molt camí per fer encara. Així doncs, trobem uns bons resultats pel que fa a l'anàlisi sintàctica del llenguatge, inicialment l'anglès i estès ja a altres llengües. En aquells casos on és necessari, també s'han resolt els problemes de morfologia (castellà, català, francès...).

Però la comprensió del llenguatge natural no es pot abordar utilitzant tant sols informació morfològica i sintàctica. Necessitem informació semàntica i pragmàtica per tal d'interpretar correctament un text o qualsevol altra manifestació del llenguatge. Moltes vegades és necessari conèixer el context per fer una interpretació adequada d'una frase. És aquí on cada cop més entren en joc les tècniques d'Intelligència Artificial i on es comencen a observar alguns resultats, sobretot en el camp del disseny d'interfícies per accedir a bases de dades.

Hem d'admetre que, ara per ara, es dificil la comprensió automàtica del llenguatge natural en tota la seva extensió. Per tant el desenvolupament d'aplicacions requereix la restricció del llenguatge: es defineixen restriccions a nivell conceptual i a nivell lingüístic.

Conscients d'aquest panorama, hem situat la nostra recerca en la problemàtica del tractament del llenguatge escrit. El treball realitzat ha consistit en estudiar un conjunt de problemes relacionats amb el tractament de textos narratius, prenent com a punt central de la recerca, la informació temporal: com s'expressa en el text, com s'interpreta, com ajuda al procés de comprensió del text i com es representa en un sistema de coneixement. Per tant, la nostra aplicació consisteix en la comprensió del llenguatge narratiu.

Els **objectius** que ens hem plantejat són:

- Dissenyar un formalisme de Representació del Coneixement capaç de contenir tot el coneixement necessari en el procés de comprensió del llenguatge: el de caràcter general, el lingüístic i el propi del domini.
- Estudiar l'expressió de la informació temporal en castellà.

- Definir mecanismes per interpretar la informació temporal dins del procés general de comprensió del text.
- Dissenyar un sistema de representació de la informació temporal integrat en un sistema de representació del coneixement.

Ha sigut necessari restringir el llenguatge, definir un domini: les notícies de diari. Però això no vol dir que les eines desenvolupades siguin vàlides tan sols en el domini escollit. El procés d'anàlisi d'un text, les tècniques de representació, el procés de raonament no depenen del domini i per tant, si volem canviar-lo, només cal modificar la base de coneixement del sistema.

COTEM és un prototipus, dissenyat a partir de la recerca realitzada, per tal d'avaluar computacionalment l'enfocament. Com a domini de treball s'ha escollit les notícies de diari. Considerem que són una bona mostra de textos narratius amb una quantitat important d'informació temporal. En particular, hem seleccionat notícies de la secció d'economia del diari "El País". Aquestes notícies fan referència, de manera objectiva, a diferents successos del món empresarial: creació, tancament o fusió d'empreses, ampliacions de capital de les empreses i resums de resultats econòmics en períodes determinats, entre altres.

La qüestió de la representació del coneixement es pot abordar, principalment, per dues vies: la lògica formal o els *frames*. Hem optat per la segona doncs permet una representació més flexible i completa dels conceptes [Minsky-75] i [FikKeh-85]. La representació mitjançant *frames* permet establir jerarquies de conceptes, definir el grau de generalitat dels descriptors, establir relacions d'herència i definir mecanismes d'inferència associats a tota la jerarquia.

Utilitzant un formalisme basat en *frames*, s'ha implementat una Base de Coneixement en la qual es representen els elements del domini (fets o estats econòmics, objectes financers, tipologia conceptual de les notícies...), el coneixement lingüístic (categorització verbal, casos semàntics, patrons temporals dels temps verbals...) i el coneixement general, amb la inclusió d'un model del temps. Les instàncies d'elements temporals i les relacions entre elles es creen i mantenen apart en el Sistema Temporal.

El Sistema Temporal dissenyat és independent del domini i consta d'una xarxa d'elements temporals de diferents tipus (punt, interval) connectada amb un calendari. Aporta la novetat, entre altres, de distingir entre elements temporals tancats i oberts segons si tenen o no, respectivament, relació amb elements de calendari. La xarxa manté els elements lligats mitjançant relacions temporals, en la línia de [Allen-81] i [Allen-83], però especialitzades en funció de si els elements a relacionar són intervals o punts. També imposa restriccions en funció de si són oberts o tancats. Aixó fa que

l'algorisme de propagació de relacions temporals sigui més eficient.

L'expressió lingüística del temps s'ha estudiat a dos nivells. A nivell de frase, proposem una classificació de les expressions temporals, els patrons E-R-S [Reichenbach-47] per temps verbals en castellà i la combinació de les expressions temporals amb els patrons verbals. A nivell de discurs, les notícies són classificades segons la seva estructura narrativa i hem definit relacions entre aquests tipus estructurals, els lexemes verbals i els temps verbals.

El prototipus implementat tracta les notícies a partir d'una representació semàntica del text. Les notícies són interpretades i la seva informació és incorporada a tot el sistema de representació (econòmic i temporal). Les regles per la interpretació de la informació temporal són independents del domini, si bé admeten restriccions del domini per tal de representar la informació de la manera més precisa possible.

Organització de la memòria

En el capítol 2 tractem de situar la nostra recerca dins del marc dels treballs realitzats en la mateixa línia o similar. Fem un repàs d'alguns sistemes de comprensió del llenguatge natural en aplicacions similars a la nostra. Presentem un parell de sistemes que només tracten expressions temporals. També s'hi descriuen els treballs més coneguts i rellevants sobre la modelització de la informació temporal, tant des del punt de vista lingüístic com des del punt de vista computacional.

El capítol 3 presenta els trets característics del nostre treball. Una petita mostra dels textos tractats dona una idea del domini seleccionat. S'han definit tipologies de les notícies a dos nivells: estructural i conceptual. Un cop presentada la tipologia de les notícies, es descriu quina és la classificació dels verbs que utilitza el sistema. Finalment es dona una visió general del funcionament de COTEM mitjançant l'exposició superficial del processament d'un exemple.

Dediquem el capítol 4 a la informació temporal: quin és el tipus d'informació temporal que apareix en el domini i com s'expressa el temps en el llenguatge. Es mostra l'estudi realitzat sobre els temps verbals del castellà, així com la tipologia definida per les expressions temporals, algunes generals i altres pròpies del domini. També es presenta la relació entre expressions temporals, temps verbals i tipus de verbs.

En el capítol 5 es defineixen els components estructurals de COTEM: la Base de Coneixement i el Sistema Temporal. Es comenten alguns detalls de la seva implementació, realitzada utilitzant un formalisme basat en *frames*.

El Sistema Temporal és presentat en detall en el capítol 6. Es descriuen els seus components, la xarxa i el calendari, justificant la decisió de la divisió entre elements temporals i elements de calendari. Al mateix temps es justifica l'existència d'elements

temporals oberts i tancats per tal de millorar l'eficiència del sistema. Es mostren els algorismes de creació, propagació i consulta de relacions temporals.

El capítol 7 fa una presentació detallada del funcionament de COTEM. De les tres etapes per les quals ha de passar un text, la primera no ha format part d'aquest treball doncs es tracta de l'objectiu d'altres treballs dins del nostre grup de recerca. En aquest capítol es descriuen les dues etapes que han sigut implementades a COTEM: el processador, que converteix la representació semàntica inicial en esquemes de l'estil dels continguts a la Base de Coneixement, i l'incorporador, que fa les comprovacions de consistència corresponents per tal d'integrar aquests esquemes a la Base de Coneixement i al Sistema Temporal. L'exposició detallada del processament de dos textos tanca la descripció del funcionament del sistema.

En el capítol 8 es presenten les conclusions del treball realitzat i les línies futures de recerca.

L'Annex relaciona un subconjunt representatiu de verbs del domini. Cada verb té una indicació del tipus al qual pertany i d'algunes peculiaritats pròpies. Al final de l'annex hi ha una relació inversa per tipus de verbs.

2. Marc bibliogràfic

Per tal de situar el nostre treball, exposem en aquest capítol a grans trets alguns sistemes dedicats a la comprensió de textos (amb o sense tractament de la informació temporal), altres sistemes especialitzats en el tractament de la informació temporal i els treballs més importants relacionats amb la modelització, comprensió i representació de la informació temporal, tant des del punt de vista lingüístic com des del punt de vista informàtic.

L'apartat 2.1 està dedicat a sistemes de comprensió de textos sense tractament de la informació temporal. Els sistemes descrits són:

- FRUMP: Sistema desenvolupat per G.F. DeJong. Es tracta d'un sistema de comprensió de textos trets directament d'un teletipus d'agència de notícies. L'objectiu de FRUMP és construir una representació del significat del text d'entrada i, mitjançant un programa de generació de llenguatge natural, produir un resum de l'estructura conceptual construïda.
- El sistema desenvolupat per H. Isahara i S. Ishizaki i que té com objectiu la comprensió de notícies sobre terrorisme escrites en japonès.
- TESS: Sistema desenvolupat per S.R. Young i P.J. Hayes. L'objectiu d'aquest sistema és la classificació i el resum de textos en anglès per dominis restringits.

Els treballs d'estudi de la modelització i representació de la informació temporal són descrits a l'apartat 2.2. S'inclou en aquest apartat tant treballs en el camp lingüístic, H. Reichenbach i N. Hornstein, com treballs en l'àrea computacional, J. Allen i altres, així com els estudis realitzats dins del marc de la traducció automàtica: EUROTRA, projecte de la Comunitat Econòmica Europea, i ROSETTA, sistema desenvolupat en el laboratori d'investigació de Philips a Eindhoven, Holanda.

A l'apartat 2.3 es presenten dos sistemes dedicats de manera única i exclusiva al tractament de la informació temporal: CHRONOS, d'en B.C. Bruce, i Time Specialist, de K. Kahn i G. Gorry.

Finalment, l'apartat 2.4 mostra alguns sistemes de comprensió de textos que tenen incorporat algun tipus de tractament de la informació temporal:

- LSP: El "Linguistic String Project" dirigit per N. Sager és un projecte ampli que ha abordat el problema de la comprensió de textos a tots els nivells. Han pres com domini de treball els informes mèdics d'un hospital.
- GROK: Sistema desenvolupat per K.K. Obermaier. L'objectiu del sistema és la comprensió de textos, per dominis restringits, de cara a la seva integració a una

base de dades. Ha sigut exemplificat utilitzant també com a textos de treball les històries clíniques.

- El grup de recerca de Toulouse dirigit per A. Borillo i M. Borillo ha estudiat la problemàtica de la comprensió del text basant-se en la Teoria de Representació del Discurs d'en H. Kamp. Proposen una implementació d'aquesta teoria a la qual han incorporat el tractament de la informació temporal.

2.1 Sistemes de comprensió de textos sense tractament d'informació temporal

FRUMP

Inicialment destaquem el sistema FRUMP descrit en detall a [DeJong-79]. Com ja hem mencionat es tracta d'un sistema de comprensió de textos trets directament d'un teletipus d'agència de notícies. L'objectiu de FRUMP és construir una representació del significat del text d'entrada i produir un resum de l'estructura conceptual construïda.

La idea bàsica és la realització d'un procés integrat de comprensió, utilitzant la combinació de criteris sintàctics, semàntics i pragmàtics. Per aquest motiu proposa els *sketchy scripts* que serveixen per descriure d'una manera estructurada el coneixement sobre les situacions del món real a tractar.

Un *sketchy script* és una agrupació estructurada de *scripts* [SchaAbe-77], que descriu participants i accions implicades en una situació. La definició d'un *sketchy script* es basa en la idea de que els fets es poden descompondre en una seqüència ordenada d'accions predefinides.

Per cada text a tractar, FRUMP ha d'identificar quin *sketchy script* cal aplicar. Un cop decidit, aquesta estructura guia el procés de comprensió. A continuació presentem el *sketchy script* per les manifestacions:

Peticció 1: Els manifestants arriben al lloc de la manifestació.

Peticció 2: Els manifestants caminen.

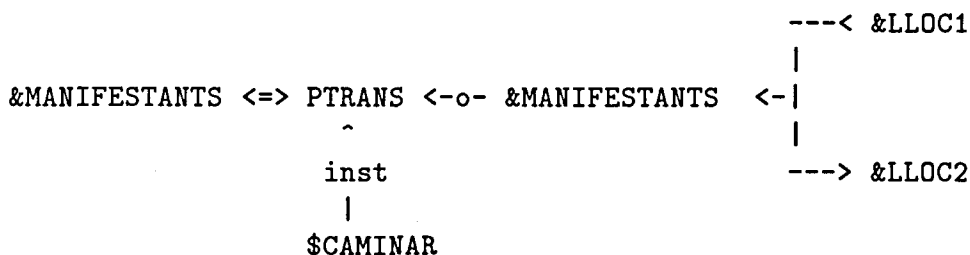
Peticció 3: La policia arriba al lloc on es troben els manifestants.

Peticció 4: Els manifestants transmeten l'objectiu de la manifestació.

Peticció 5: Els manifestants ataquen l'objectiu de la manifestació.

Peticció 6: Els manifestants ataquen la policia.

FRUMP representa aquestes peticions en termes de l'estructura conceptual d'en Schank [Schank-72]. Per exemple, la petició 2 es representa:



Cada petició té (1) restriccions semàntiques sobre les variables, (2) restriccions entre les variables d'un *script* i (3) relacions causals amb altres peticions i referències a altres *sketchy scripts*.

El funcionament de FRUMP es basa en dos mòduls: el *predictor* i el *substantiator*. El primer fa una predicció de restriccions conceptuals sobre el que pot succeir en el text, per realitzar la seva tasca aquest mòdul fa servir els *sketchy scripts*. El segon intenta verificar aquestes prediccions, directament del text o mitjançant inferències.

El primer plantejament del nostre treball, la comprensió de notícies, va consistir en intentar aplicar aquesta teoria al domini seleccionat per nosaltres. Fet un estudi dels textos, es va observar que els trets fonamentals de les notícies es podien extreure per altres mecanismes, mentre que els *sketchy scripts* no resultaven massa adequats.

Els textos del nostre domini no presenten restriccions sobre l'ordre en el qual apareixen les accions, per tant és fa difícil definir *sketchy scripts*. En canvi sí es pot definir el conjunt de conceptes que poden aparèixer segons la notícia. Això permet realitzar una organització conceptual del domini. D'altra banda, el nostre interès pel tractament de la informació temporal ens ha permès detectar el seu ajut en el procés de comprensió de la notícia. Hem considerat més interessant i escaient dissenyar un mecanisme d'interpretació de la notícia, basat en una doble caracterització del text (conceptual i estructural), en lloc d'intentar ampliar els *sketchy scripts* amb restriccions temporals.

Tractament de notícies en japonès

Seguint en el domini de les notícies, el treball descrit a [IsaIshi-86] és també un sistema de comprensió de textos, notícies en japonès sobre terrorisme. A diferència de FRUMP, el que descrivim aquí és un exemple de sistema on es realitza seqüencialment una anàlisi sintàctica, una semàntica i una contextual del text per tal d'obtenir la seva representació contextual. Aquestes representacions són una mena d'escenaris on es descriu la situació mitjançant les accions, els participants i les relacions entre ells. La figura 2.1 mostra un exemple d'estructura contextual.

L'analitzador sintàctic obté del text unes estructures sintàctiques preliminars. L'analitzador semàntic tracta de modificar tant les relacions internes dels grups nominals

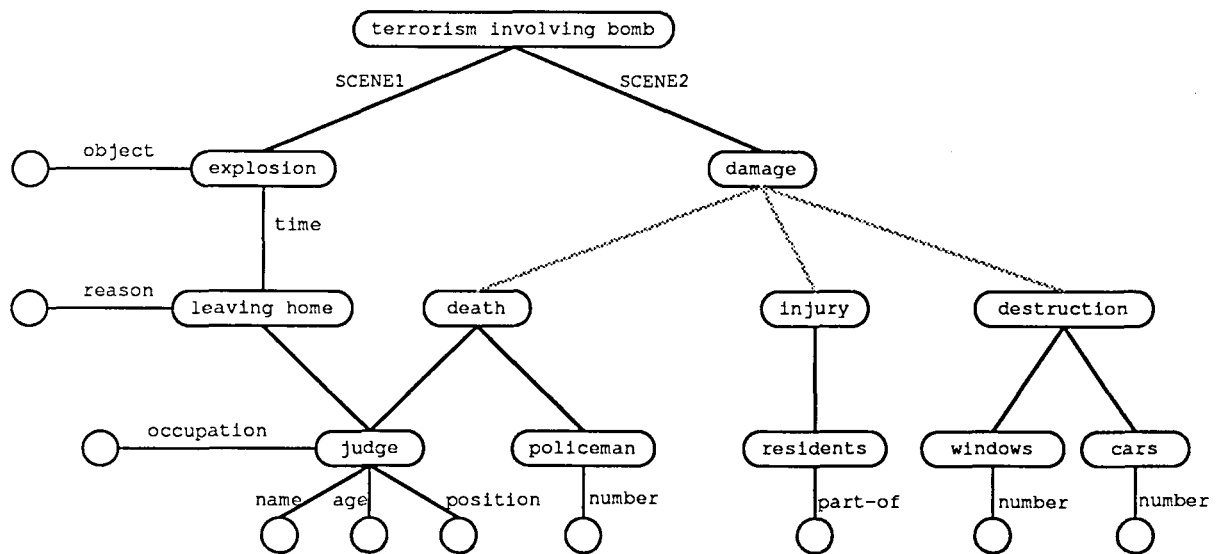


Figura 2.1

com les relacions de casos entre els verbs i els grups nominals. Finalment, l'analitzador contextual obté la informació contextual a partir de les estructures semàntiques generades per l'analitzador semàntic. El sistema conté coneixement sobre el terrorisme i tracta d'incorporar a la seva estructura contextual la informació sobre els fets descrits en els textos.

TESS

Un altre sistema interessant és TESS, sistema per classificar i resumir textos en anglès per dominis restringits. Una versió d'aquest sistema adaptada a l'anàlisi de télexs bancaris està comentada a [YouHa-85].

El funcionament de TESS consta de tres fases: (1) classificació del télex segons una jerarquia predefinida, (2) extracció de la informació relevant a partir de les prediccions que faci la classificació inicial i (3) inferència de la informació addicional necessària que no està explícita en el télex però que es pot deduir de les dades extretes a la fase anterior.

Cadascuna de les categories de la jeraquia conceptual té associat un *expectation frame*, els descriptors del qual especifiquen la informació a extreure dels textos corresponents a la categoria en qüestió. Es pot considerar que els descriptors proporcionen expectatives sobre la informació que pot aparèixer en el text. L'*expectation frame* associat a les tranferències bancàries és:

Descriptors	Valors omplidors	Obligatorietat / Valors per defecte
Remitent	Localització Nom Banc	Obligatori
Receptor	Localització Banc	Obligatori
Data	Dia Mes Any	Obligatori
Codi Seguretat		Obligatori
Import Petició	Moneda Import	Obligatori
Compte Crèdit	Identificador Compte	Obligatori
Lloc del Compte Crèdit	Ciutat del Banc	Banc receptor
Compte Dèbit	Identificador Compte	Remitent
Lloc del Compe Dèbit	Ciutat del Banc	Banc remitent
Motiu	Motiu	Opcional

El funcionament de TESS té una petita semblança amb el de COTEM. El primer pas per tots dos sistemes consisteix en una classificació del text. Però mentre TESS situa el text dins una jerarquia conceptual, COTEM classifica el text des d'un punt de vista conceptual i un altre estructural. Els processos per extreure la informació ja són totalment diferents sobretot tenint en compte el tractament de la informació temporal, inexistent a TESS.

2.2 Modelització i representació de la informació temporal

Sobre la modelització de la informació temporal treballen tant lingüistes com informàtics, de manera totalment independent la majoria de vegades. COTEM evita aquesta separació d'interessos combinant el coneixement lingüístic amb una proposta de representació computacional. Per aquest motiu, hem consultat els treballs realitzats en ambdues àrees.

Enfocament lingüístic

Pel que fa a l'estudi de l'expressió lingüística del temps cal situar-se en el punt de partida de la majoria de treballs: la teoria dels tres identificadors, E-R-S, d'en H. Reichenbach [Reichenbach-47]. Aquesta teoria proposa l'existència de tres identificadors associablès a cada frase a partir del temps verbal: S (*speech moment*), E (*event moment*) i R (*reference moment*). Entre aquests identificadors hi ha dues

relacions possibles: anterioritat (-) i anterioritat-o-contemporanis (,). En el capítol 3 tornem a parlar d'aquesta teoria i mostrem l'ús que COTEM en fa d'ella.

Aquesta teoria ha sigut utilitzada i ampliada per molts investigadors posteriors. És el cas d'en N. Hornstein [Hornstein-77] que desenvolupa una teoria per combinar les estructures E-R-S, definides pels temps verbals, amb els adverbis temporals per tal de donar una interpretació temporal global de la frase. També tracta algunes connectives temporals com *when, after, before, until, since*.

Hornstein anomena estructura temporal bàsica a la tripleta E-R-S definida a partir del temps verbal, mentre que considera com estructura temporal derivada la que resulta de combinar l'estructura temporal bàsica amb els adverbis temporals. Per exemple:

- El model definit pel *Simple Present* és S,R,E. Aquesta és l'estructura temporal bàsica.
- Donada la frase *John comes home tomorrow*, l'estructura temporal derivada és S..R,E amb *tomorrow* associat a R,E.

Per tal de considerar si una estructura temporal derivada és vàlida, Hornstein proposa algunes regles. Al combinar l'estructura E-R-S del temps verbal amb un adverbi temporal, pot ser que hi hagi un canvi d'ordre entre els identificadors i/o un canvi de les relacions existents entre ells. Hornstein no accepta una estructura temporal derivada si (1) l'ordre línia E-R-S de l'estructura bàsica no es manté i/o si (2) l'estructura derivada permet alguna possibilitat d'igualtat (contemporanis) entre els descriptors que l'estructura bàsica no tenia.

COTEM també realitza un procés de combinació més ampli doncs accepta expressions temporals en general, expressions que han sigut classificades com exposen en el capítol 3.

Ambdues teories lingüístiques consideren que aquests identificadors són punts i per tant les relacions possibles queden reduïdes a les dues mencionades.

En els treballs del projecte europeu de disseny d'un sistema de traducció automàtica multilingua, EUROTRA, [ArToJa-85] també s'ha seguit la teoria dels E-R-S, però considerant que són intervals i no punts. S'han basat en el treball d'en B.C. Bruce que descrivim a l'apartat 2.3. Per exemple, la frase *John died in 1983* és interpretada des del punt de vista temporal com:

E during R & (in 1983) R & Rbefore S

També han fet algunes consideracions sobre com afecta l'aspecte verbal a les estructures E-R-S. Afegeixen un quart identificador (Q) per tal de tractar l'aspecte.

En publicacions posteriors, [Eynde-87] i [Eynde-88], presenten un model temporal del discurs.

Un altre sistema de traducció automàtica on també s'han estudiat les qüestions temporals és el ROSETTA [Appelo-86], aquest sistema es basa en el Principi de Composicionalitat de Montague. També dóna com referència principal els treballs d'en B.C. Bruce i també es tracta parcialment la qüestió de l'aspecte verbal. El mètode de traducció en aquest sistema s'anomena *isomorphic grammar method* i, pel que fa a l'informació temporal, consisteix en definir per cada llengua a la que es vol traduir:

- les expressions temporals que expressen intervals
- les regles sintàctiques que indiquen com es combinen les expressions temporals de cara a expressar relacions entre intervals.

Per cada expressió a la gramàtica d'una llengua estableixen com a mínim una expressió dins la gramàtica de les altres llengües. Per les regles sintàctiques fan el mateix.

Consideren tant el temps verbal com les expressions temporals: adverbis, grups preposicionals i conjuncions temporals. El model temporal consta d'intervals als quals s'associen propietats com, per exemple, la longitud o durada. Consideren que hi ha una relació déictica entre R i S mentre, entre R i E, existeix una relació aspectual. Les propostes d'interpretació són similars a les del projecte EUROTRA, però fent més èmfasi en la qüestió de l'aspecte i donant propostes concretes de mecanismes de traducció.

Enfocament computacional

Des del punt de vista computacional, els treballs més coneguts són el model d'en J. Allen [Allen-83] i la lògica basada en punts d'en D. McDermott [McDermott-82]. Descrivim el treball més ampliament seguit, és a dir, la proposta d'en J. Allen: una lògica temporal basada en intervals combinada amb un algorisme de raonament temporal, basat en la propagació de relacions.

Defineix com a relacions temporals possibles: *equal, before, after, meets, met-by, overlaps, overlapped-by, during, contains, starts, started-by, finishes i finished-by*. Entre dos elements temporals poden haver-hi varies relacions temporals quan no es coneix la relació exacta. Una taula de transitivitat de relacions indica quina és la relació composició al combinar dues relacions qualsevol. Per exemple: donades les relacions a R1 b i b R2 c, essent R1 *overlaps* i R2 *during*, es dedueix a R3 c, essent R3 = (*overlaps / during / starts*).

COTEM es basa en la proposta d'en J. Allen si bé l'amplia per tal de donar entitat pròpia als punts, prou importants en el domini com per tenir un tractament diferenciat.

En aquesta línia també van els treballs exposats a [Vilain-82] i [VilKau-86]. M.B. Vilain justifica el tractament dels punts en que els intervals no són l'únic mecanisme amb el qual els éssers humans interpreten el temps. També recorda que els primers sistemes de representació temporal (justament els que es descriuen a l'apartat 2.3) tenen el punt com element bàsic de representació.

Un treball força interessant és el descrit a [Yip-85]. Hi trobem una combinació de la teoria lingüística d'en N. Hornstein i la teoria computacional d'en J. Allen. Consegueix combinar les restriccions lingüístiques del primer amb el model temporal del segon, demostrant que és un model vàlid també des del punt de vista lingüístic.

Amb posterioritat, [AllHay-85] i [AllHay-87], s'ha proposat la mateixa lògica temporal utilitzant una única relació temporal: *meets*.

Altres treballs sobre representació temporal són els descrits a [Kandrashina-83] i [Ladkin-86]. Ambdós consideren la qüestió dels intervals temporals en la línia de J. Allen, i ambdós continuen considerant el punt com un interval degenerat (punt d'inici = punt de final). Però tenen la particularitat de ser de les poques referències que existeixen sobre cadenes: seqüències d'intervals no necessàriament consecutius. A [LeDoFor-86] s'exposa un treball dedicat exclusivament a les cadenes temporals. COTEM no tracta actualment aquest tipus d'element temporal, però té previst el seu tractament en ampliacions posteriors del sistema.

2.3 Sistemes especialitzats en el tractament de la informació temporal

CHRONOS

Aquest sistema presenta un model per resoldre referències temporals i fa una proposta d'implementació. La descripció en detall del model, amb el tractament dels temps verbals i altres relacions temporals d'ordre, més una mostra del sistema es troba a [Bruce-72]. L'ampliació del sistema per tal de tractar expressions temporals que contenen punts temporals explícits es presenta a [Bruce-73].

En B.C. Bruce analitza els temps verbals, les relacions temporals i altres referències al temps en el llenguatge. El tractament temporal es fa a nivell de frase. Defineix *time* com un conjunt d'elements anomenats *time points* i defineix com *time-system* la parella formada per aquest conjunt i la relació \leq , que parcialment ordena el temps. Altres conceptes que utilitza són *time-segment*, *duration*, *time-segment relation*, *tense*, *reference time* i *tense marker*.

CHRONOS tracta punts i intervals, permet més d'un referent temporal i utilitza les estructures E-R-S considerant que són intervals. Per tant apareixen relacions com solapa, conté, durant... que arriba a definir en termes de les dues relacions entre punts.

Els elements temporals són representats mitjançant cinc descriptors (any, mes, dia, hora, minut) que tenen valors per defecte. Per exemple, el 15 de gener de 1989 es representa en aquest sistema com (1989 1 15 0 0) mentre que les 10.30 hores del 20 d'agost de 1988 es representa per (1988 8 20 10 30). Aquest tipus de llista també serveix per representar durades, per exemple (3 1 0 0 0) s'interpretarà com tres anys i un mes.

CHRONOS es capaç de seleccionar predicats (verbs) i arguments (noms) a partir de les frases, i de fer deduccions temporals. Segons l'autor, sembla que el sistema es podria integrar en un sistema més general de pregunta-resposta. Tanmateix, no fa deduccions basant-se en el significat dels predicats. A l'igual que Time Specialist, que descrivim a continuació, només entén preguntes sobre qüestions temporals, no coneix res sobre el domini en curs.

Time Specialist

Sistema presentat en detall a [KahGor-77]. Es tracta d'un sistema de tractament de la informació temporal que es pot integrar a un sistema de caràcter general. Es basa en l'hipòtesi de que el coneixement sobre el temps es pot representar com un conjunt de procediments de *problem-solving* que col·lectivament anomenen Time Specialist.

El sistema tracta d'interpretar *temporal specifications*, frases que especifiquen parcialment el temps d'un o més fets. Es considera que en una *temporal specification* hi ha alguna relació entre dos fets, com a mínim, i cada fet es considera un punt en el temps. El sistema posseeix procediments per comparar expressions temporals de diferents tipus, per fer traduccions entre elles i per combinar expressions temporals fent inferències. Les úniques relacions temporals que contempla són les definides entre punts: abans, després o igual.

Time Specialist té diferents formes d'organitzar els fets a la memòria: per dates (els fets són situats de manera cronològica sobre la *date line*), per *special reference events* (fets dels quals es coneix el moment temporal i que serveixen per deduir el moment temporal d'altres fets que presenten una relació temporal amb aquests especials) i per *before/after chains* (en el cas que es pugui establir seqüències de fets, el sistema crea cadenes amb ells). Es pot aplicar més d'un d'aquests criteris alhora per tal d'interpretar una nova entrada al sistema.

Aquest sistema també considera la problemàtica del manteniment de la consistència a la Base de Dades.

2.4 Sistemes de comprensió de textos amb tractament d'informació temporal

Linguistic String Project

Com ja hem dit, el "Linguistic String Project" (LSP) que dirigeix N. Sager aborda el problema de la comprensió de textos a tots els nivells. És molt interessant i complet el sistema d'anàlisi sintàctica que han desenvolupat [Sager-81]. Aquest sistema ha servit d'orientació en altres treballs del nostre grup de recerca [Rodríguez-89] i que COTEM ja no ha abordat. El LSP ha anat més enllà de l'anàlisi sintàctica, dedicant-se també a l'estudi de la comprensió de textos per dominis restringits. En particular, han pres com domini de treball els informes mèdics d'un hospital [SaFriLy-87].

En el LSP els peculiars textos d'entrada són analitzats i representats utilitzant els *information formats*, una estructura pròpia d'aquest sistema relativament similar a un *frame*.

Aquest sistema contempla el tractament de la informació temporal que apareix a les històries clíniques. El temps és un dels possibles modificadors en els *information formats*. La implementació realitzada és una solució molt particular del domini doncs es tracta d'un mecanisme senzill capaç de representar en forma de graf la informació temporal dels textos, que no és massa complexa. Bàsicament utilitzen les distàncies temporals entre dos fets més que els moments temporals associats a aquests.

L'*information format* que correspon a la frase *Most recent follow-up visit in ophthalmology clinic was 3/1/85* és

FORMAT1	(GENERAL MEDICAL MANAGEMENT)
INST	ophthalmology clinic (in)
PT	
VERB-MD	follow-up visit
{**}	
VERB	was
{TENSE}	[PAST]
*EVENT-TIME	
TREL1	
NUM	
UNIT	
TREL2	most recent
REF-PT	visit

*EVENT-TIME

TREL1
NUM
UNIT
TREL2
REF-PT 3/1/85

Com exemple d'un altre tipus d'expressió, el grup preposicional *for 3 days prior to admission* es representa com:

EVENT-TIME

TREL1 for
NUM 3
UNIT days
TREL2 prior to
REF-PT admission

A partir d'aquestes estructures, el sistema dedueix la durada dels fets i aixó és el que es representa com un arc en el graf temporal.

GROK

GROK és un sistema de comprensió de textos per dominis restringits que també ha sigut exemplificat utilitzant històries clíniques, [Obermeier-85a] i [Obermeier-85b]. El programa analitza el text utilitzant coneixement lingüístic i extra-lingüístic (semàntic i conceptual). Les estructures resultants de l'anàlisi són traduïdes a representacions de fets, sobre les quals es realitzen inferències tant temporals com del domini. I finalment s'actualitza la Base de Dades amb la informació relevant sobre el malalt.

El tractament de la informació temporal que realitza GROK es basa en la detecció en el text dels *key events* propis del domini. Aquests elements serveixen per construir l'estructura temporal del fet. Exemples de *key events* en domini indicat són: *at admission* o *at surgery*.

Posa l'èmfasi en els adverbis i considera el temps verbal com la característica temporal menys relevant. L'ordre de les frases és important doncs considera que el temps avança a mesura que es progressa en el text, si no es diu cap cosa en contra. La referència temporal per les oracions subordinades es considera la mateixa que per les principals.

Grup de Toulouse (IRIT i Equipe de Recherche en Syntaxe et Sémantique)

Sota la direcció de M. Borillo i A. Borillo, el grup de recerca de Toulouse, format per lingüistes i informàtics, ha estudiat també la problemàtica de la comprensió del text [BoBoBras-89]. La semàntica del temps, tant a nivell de frase com de discurs, ha sigut abordada des del punt de vista teòric i des del punt de vista de la implementació. Utilitzen la teoria DRT (*Discours Representation Theory*) d'en H. Kamp, [Kamp-81] i [KamRoh-83].

El domini lingüístic ha sigut reduït a verbs de localització, verbs de desplaçament, adverbis de lloc i adverbis temporals, tant deíctics com anafòrics. Es tracten només frases declaratives i el temps verbal és sempre el passat (*Passé Simple, Passé Composé, Imparfait, Plus Que Parfait*). L'objectiu és extreure la informació temporal del text a partir de l'aspecte i temps verbals i de la semàntica dels adverbis temporals.

El seu formalisme de representació es basa en arbres de tres branques [Colmerauer-79] als quals s'afegeix una component temporal.

Frase Temporalitzada \rightarrow component temporal + Frase (arbre 3 branques)

La component temporal consta de quatre elements:

- TV: Temps verbal.
- CV: Categoria aspectual del verb
- ET: Entitat temporal associada a la frase
- Component adverbial:
 - Traits_a: Trets semàntics
 - Fla: Forma lògica de l'adverbi
 - Cta: Constant temporal associada

Definida la representació a nivell de frase, passen a l'estudi de les estructures discursives: la construcció de la DRS (*Discours Representation Structure*). Tenen regles locals, per tractar els referents temporals i els no temporals a nivell de frase, i regles a nivell de discurs, per construir les condicions temporals. Algunes regles a nivell de discurs són:

* El darrer succés incorporat a la DRS es marca com a punt de referència R.

- Si la nova entitat introduïda és un succés $e(i)$ i essent el temps verbal *Passé Simple*:

$R < e(i) ; e(i) < n$ (n és el moment actual)

A partir d'ara $e(i)$ serà el nou R.

- Si la nova entitat introduïda és un estat $s(i)$ i essent el temps verbal *Imparfait*:

$R \subset s(i)$

- * Per frases completades amb adverbis temporals: combinar les informacions aportades per l'adverbi i el verb.

La representació de *A trois heures, le robot entra dans la navette. Puis, il traverse le sas* presenta els següents elements:

x(1)	e(1)	l(1)	ct(1)	l(2)
ct(1) := loc(3,h)				
robot(x(1))				
navette(l(1))				
e(1): entrer_dans(x(1),l(1))				

* ct(1) \equiv e(1)

* e(1) < n

sas(l(2))

e(2) : traverser(x(1),l(2))

* e(1) < e(2)

* e(2) < n

Essent e(i) l'identificador temporal del succés i, l(j) l'identificador del lloc j, x(k) l'identificador de l'individu k i ct(n) la n-èsima constant temporal.

Les línies marcades amb * són resultat d'aplicar les condicions temporals a nivell de discurs.

També s'han estudiat alguns components espaials i com poden modificar l'estructura temporal.

2.5 El sistema COTEM

El nostre sistema presenta trets d'alguns dels sistemes descrits en els apartats anteriors.

COTEM conté una classificació conceptual del coneixement del domini que li permet definir una tipologia també conceptual dels textos. Aquesta classificació es complementa amb una tipologia estructural dels mateixos textos. El sistema realitzarà un procés de classificació a l'igual que realitza TESS.

Considerem que el processament d'un text de cara a la seva comprensió requereix l'utilització de coneixement morfològic, sintàctic, semàntic i de domini. El sistema

descriu a [IsaIshi-86] aplica aquests coneixements de manera seqüencial, mentre que COTEM es planteja la combinació de tots ells. En particular, les tipologies mencionades anteriorment intervenen de manera fonamental en el procés de comprensió del text.

A més, COTEM aporta un model de comprensió i representació del temps. S'han combinat teories lingüístiques, sobre els temps verbals i les expressions temporals, amb el model d'en J. Allen per tal de definir un model temporal propi. Els capítols 3 i 4 mostren de manera concreta com s'ha fet aquesta combinació.

La nostra recerca aporta la integració del processament del llenguatge natural, amb la representació del coneixement i amb el tractament de la informació temporal en un sistema de comprensió de textos narratius en castellà.

3. Caracterització de l'aplicació: domini i exemple

La font d'informació del sistema són les notícies d'agència que apareixen a les pàgines de la secció d'economia del diari El País. Aquestes notícies fan referència d'una manera objectiva a: la creació, tancament o fusió d'empreses, la compra d'una empresa totalment o parcialment per part d'una altra, les inversions i les ampliacions de capital realitzades per una empresa, els beneficis i pèrdues en el resum anual de les empreses, etc.

La mostra recollida consta actualment de 128 notícies, agafades en diferents dies de la setmana, diferents dies del mes i diferents mesos. Fins i tot, a la mostra inicial s'han anat incorporant notícies i per tant n'hi ha d'anys diferents. Aquesta ampliació s'ha fet per tal d'aconseguir una mostra més completa sobretot d'aquelles notícies menys freqüents. Normalment, les notícies no estan signades per cap persona, són notícies d'agència que indiquen simplement la ciutat i, com a molt, van signades per EL PAIS. No presenten opinions sobre els fets descrits, contenen tan sols l'exposició de dades mitjançant frases declaratives.

S'ha realitzat una anàlisi de la mostra amb la finalitat de modelitzar el coneixement sobre el domini, coneixement fonamental per la comprensió dels textos, i de caracteritzar les notícies. Com a resultat d'aquesta anàlisi hem determinat, per una banda, la Base de Coneixement sobre el món empresarial i, per una altra, una tipologia de notícies que descrivim a continuació.

3.1 Tipologia de les notícies

Dos han sigut els criteris aplicats a l'hora de classificar les notícies: contingut de les notícies, criteri de domini, i estructura del text, criteri del tipus d'aplicació.

Classificació conceptual

Segons la informació central de les notícies, aquestes han sigut classificades d'acord a tres tipus conceptuals.

En el tipus 1 s'inclouen totes les notícies referents a fusions d'empreses, integracions d'una empresa en una altra, compres d'empreses o de participacions, ofertes de venda o d'adquisició d'accions, i les operacions (compra, oferta venda, comanda) sobre objectes que no són empreses com, per exemple, hotels, avions, etc.

Totes les notícies del tipus 1 tenen en comú la presència d'un agent principal de tipus "persona", concepte genèric que conté tant la persona física com la persona jurídica (corporació, fundació, empresa, municipi...). Aquests textos són subclassificats en funció del tema tractat que pot ser un objecte financer (1.1) en el cas de les operacions financeres o un objecte físic (1.2) en el cas de transaccions comercials.

Considerem objectes financers de dos tipus: l'empresa com a tal entitat i les participacions, percentatge d'accions en aquelles empreses de propietat distribuïda entre accionistes. Segons el tipus d'objecte financer, les notícies del tipus 1.1 es tornen a subclassificar. Les que tenen com a tema una empresa (1.1.1) presenten una restricció addicional sobre l'agent, ja no és qualsevol tipus de "persona" sinó que es tracta també d'una empresa. Són operacions de fusió entre empreses (1.1.1.1), per absorció o per entente, o d'integració d'una empresa a una altra (1.1.1.2) o de compra d'una empresa per part d'una altra (1.1.1.3). D'aquest darrer subtipus hi ha un exemple a la figura 3.1. Si el tema de la notícia són participacions d'una empresa (1.1.2), ens pot informar sobre una oferta de compra d'accions (1.1.2.1), pública o privada, o sobre la venda d'un percentatge d'accions (1.1.2.2) que pot arribar a ser del 100%.

Entre les transaccions comercials distingim tres subtipus: les comandes (1.2.1) d'objectes que cal construir (avions, vaixells, edificis...), les compres d'objectes físics ja existents (1.2.2) com pot ser un hotel i les ofertes de venda d'objectes físics existents (1.2.3) com per exemple una fàbrica. A la figura 3.1 es presenta una notícia que informa sobre la compra d'un hotel.

EL PAIS, sábado 18 de enero de 1986

Hunter Douglas amplía su grupo con la compra de dos empresas

Madrid
Hunter Douglas ha ampliado su grupo de empresas con la compra de dos compañías, Filtrasol, SA, de Francia, y Stilsound Holding Limited, del Reino Unido.

Filtrasol ocupa a una plantilla de 380 empleados y posee un volumen de ventas de unos 4.200 millones de pesetas anuales. Stilsound Holding, por su parte, emplea a 160 personas. Sus ventas anuales se sitúan en torno a los 1.300 millones de pesetas.

El grupo Hunter Douglas está especializado en la fabricación de productos de aluminio para la decoración y la arquitectura, y está representado en España a través de su filial Hunter Douglas España, SA.

El grupo británico Bass adquiere el hotel Villamagna

Madrid
El grupo Bass, uno de los mayores *holdings* británicos y segunda organización hotelera del Reino Unido, ha adquirido el hotel Villamagna de Madrid por 3.100 millones de pesetas.

La operación, que ha sido aprobada por el Gobierno español, se ha realizado a través de Bass European Holdings tras un largo proceso de negociaciones con la empresa vendedora —Nerja, SA—, y en competencia con diversas firmas internacionales.

Junto a una inversión de 230 millones de pesetas destinada a modernizar su equipamiento, Bass se propone mejorar los servicios del establecimiento madrileño cara a la expansión del grupo en el sector de hoteles de lujo.

Figura 3.1

L'esquema de la classificació de les notícies de tipus 1 és el següent:

1 CANVI-DE-PROPIETAT

Agent: persona

1.1 OPERACIÓ-FINANCERA

Tema: objecte-financer

1.1.1 TOT

Tema: empresa

Agent: empresa

1.1.1.1 FUSIÓ

absorció

entente

1.1.1.2 INTEGRACIÓ

1.1.1.3 COMPRA

1.1.2 ACCIONS

Tema: participacions

1.1.2.1 OFERTA

OPA

privada

1.1.2.2 VENDA

total

parcial

1.2 TRANSACCIÓ-COMERCIAL

Tema: objecte-físic

1.2.1 COMANDA

Objecte-físic-a-construir

1.2.2 COMPRA-DIRECTA

Objecte-físic-existent

1.2.3 OFERTA-VENDA

Objecte-físic-existent

Qualifiquem de notícies de tipus 2 totes aquelles que relacionen resums econòmics de les empreses. Són notícies que declaren els beneficis, les inversions, les facturacions... d'una empresa en un període de temps determinat. També es poden donar resums econòmics per altres tipus de persones jurídiques, però mai l'agent d'aquest tipus de notícies és una persona física. A la figura 3.2 es mostra una notícia d'aquest tipus.

Los beneficios de Eurobuilding se incrementaron un 55% en 1986

R. V., Madrid

Los beneficios de Eurobuilding se han incrementado en el ejercicio pasado en un 55%, lo que permite a esta sociedad repartir un dividendo del 6%, superando en dos puntos al repartido en 1985. Los ingresos de explotación han aumentado en un 54%, alcanzando los 303 millones de pesetas.

Asimismo, la sociedad va a realizar, posiblemente en abril, una ampliación de capital a la par, en la proporción de una acción nueva por cada 12 antiguas, con el objeto de acudir a la última ampliación de capital de Eurobuilding Internacional, promotora de Eurobuilding Caracas.

Corning Glass y Telefónica crean una empresa de fibra óptica

Madrid

Corning Glass y Telefónica alcanzaron ayer en Nueva York un acuerdo para crear una empresa que fabricará y comercializará fibra óptica en España. Dicho acuerdo está sujeto a la aprobación de los respectivos Gobiernos y consejos de administración.

Corning Glass es, junto con ATT, el principal fabricante mundial de fibra óptica. La firma cuenta con 58 plantas, emplea a 27.000 personas y factura 1.737 millones de dólares (más de 270.000 millones de pesetas). Corning Glass contará con el 65% del capital en la empresa y Telefónica ostentará el 35% restante. La inversión será de 20 millones de dólares.

Figura 3.2

2 AVALUACIÓ-ACTIVITAT-EMPRESA

Agent: persona-jurídica

gastos
inversiones
facturaciones
beneficios
pérdidas
pago dividendos
ventas

...

Al tipus 3 corresponen les notícies referents a la creació d'una nova empresa (3.1), al tancament d'una empresa (3.2) o al procés d'una ampliació de capital (3.3). Es tracta del que anomenem fets propis de la vida de l'empresa. En general, l'agent és una persona jurídica doncs, per exemple, es pot donar el cas de la creació d'una empresa per part d'una entitat que no sigui també una empresa. A la figura 3.2 presentem una notícia sobre creació d'una empresa.

3 FETS-PROPIS-VIDA-EMPRESA

Agent: persona-jurídica

3.1 CREACIÓ

per entente

directa

per fusió

3.2 AMPLIACIÓ-CAPITAL

3.3 TANCAMENT

propi

per absorció

Classificació estructural

Des del punt de vista estructural, hi ha dos tipus de notícies: el TEXT-FET i el TEXT-ESTAT, segons la seva descomposició en *frames* verbals.

L'estructura d'un "text-fet" és:

Notícia principal

- 1) Titular
- 2) Introducció
- 3) Ampliació de la informació (sobre el fet o sobre els participants)

Notícies addicionals

Informacions de tipus resum econòmic sobre els participants i/o previsions d'altres fets

Les notícies dels grups 1 i 3 de la tipologia anterior presenten aquesta estructura. A la figura 3.3 es mostra la descomposició d'una notícia d'aquest tipus.

L'estructura d'un "text-estat" és:

Descripció de resums

- 1) Titular
- 2) Descripcions d'estats (resums econòmics)
- 3) Ampliació de la informació sobre l'agent

Notícies addicionals

Informacions sobre fets econòmics (no resums) passats o previstos

Les notícies del grup 2 responen a aquest segon tipus d'estructura. La figura 3.4 presenta un exemple de descomposició estructural.

Hunter Douglas amplía su grupo con la compra de dos empresas

Titular

Madrid
Hunter Douglas ha ampliado su grupo de empresas con la compra de dos compañías, Filtrasol, SA, de Francia, y Stilsound Holding Limited, del Reino Unido.

Introducció

Filtrasol ocupa a una plantilla de 380 empleados

Ampliació informació

y posee un volumen de ventas de unos 4.200 millones de pesetas anuales.

Ampliació informació

Stilsound Holding, por su parte, emplea a 160 personas.

Ampliació informació

Sus ventas anuales se sitúan en torno a los 1.300 millones de pesetas.

Ampliació informació

El grupo Hunter Douglas está especializado en la fabricación de productos de aluminio para la decoración y la arquitectura,

Ampliació informació

y está representado en España a través de su filial Hunter Douglas España, SA.

Ampliació informació

Figura 3.3

El primer *frame* verbal (el titular) s'utilitza per fer la classificació estructural de la notícia i, sempre que sigui possible, per la detecció del tipus conceptual de la mateixa. Les notícies de tipus 1 són classificables fins a un cert nivell amb el titular, però per fer-ne una definició total cal esperar a processar com a mínim la introducció.

Un cop detectat a quina estructura respon el text, ja tenim un fil conductor del procés d'interpretació de la notícia. La informació temporal és relevant de cara a definir cadascuna de les parts estructurals doncs tenen un temps verbal propi. L'estructura detectada al processar el titular genera unes expectatives que es refinen segons els temps verbals i els tipus dels verbs que anem trobant, també el tipus conceptual de la notícia pot servir per situar estructuralment alguna part de la notícia. La classificació semàntica dels verbs és fonamental per dissenyar el processament de la notícia.

Los beneficios de Eurobuilding se incrementaron un 55% en 1986

Titular

R. V., Madrid
Los beneficios de Eurobuilding se han incrementado en el ejercicio pasado en un 55%,

Descripció estat

lo que permite a esta sociedad repartir un dividendo del 6%,

Descripció estat

superando en dos puntos al repartido en 1985.

Descripció estat

Los ingresos de explotación han aumentado en un 54%,

Descripció estat

alcanzando los 303 millones de pesetas.

Descripció estat

Asimismo, la sociedad va a realizar, posiblemente en abril, una ampliación de capital a la par, en la proporción de una acción nueva por cada 12 antiguas, con el objeto de acudir a la última ampliación de capital de Eurobuilding Internacional, promotora de Eurobuilding Caracas.

Notícia adicional

Notícia adicional

Figura 3.4

3.2 Tipologia dels verbs

Els verbs han sigut classificats segons la funció que tenen en el domini. Distingim tres tipus: ACCIÓ, DESCRIPCIÓ i OPERADOR.

Són del tipus ACCIÓ tots els verbs que defineixen els successos del domini de manera directa (gestió) o de manera parcial (gestió mínima), els que defineixen accions associades a aquests successos (administrativa, resultativa) i tots els verbs d'ajut propis del tipus de discurs, una notícia de diari (discursiva, destructiva).

Són del tipus DESCRIPCIÓ tots els verbs que descriuen alguna cosa respecte un agent, normalment empreses, i tots aquells que descriuen o qualifiquen algun succés. Els verbs d'aquest grup tenen com a característica principal que sempre els trobem en *Presente* o en *Pretérito Imperfecto* d'Indicatiu.

Són del tipus OPERADOR aquells verbs que no diuen res (buit) i cal veure l'acció que acompanyen, de fet serveixen per donar el seu temps verbal a l'acció acompanyada.

També pertanyen a aquest grup els verbs que modifiquen el sentit d'allò que acompanyen (qualificador), donen un caire de futur o de previsió o incoatiu. Finalment, s'inclou també en aquest tipus unes nominalitzacions peculiars del domini que són anomenades operador-resum-gestió i que porten implícita una comparació.

La classificació completa es presenta a continuació. Cada tipus de verb porta uns codis numèrics que indiquen en quins tipus de notícia del domini apareix.

ACCIÓ

ADMINISTRATIVA 1, 3

DESTRUCTIVA 1, 2, 3

DISCURSIVA 1, 2, 3

GESTIÓ

AMPLIACIÓ-DE-CAPITAL 3.2

COMANDA 1.2.1

COMPRA-VENDA 1.1.1.3, 1.1.2, 1.2.2, 1.2.3

CREACIÓ 3.1, 1.1.1.1

FUSIÓ 1.1.1.1, 3.1, 3.3

INTEGRACIÓ 1.1.1.2

RESUM-GESTIÓ 2

TANCAMENT 3.3, 1.1.1.1, 3.1

GESTIÓ-MÍNIMA

RESULTATIVA 1, 3.1, 3.2

DESCRIPCIÓ

ACTIVITAT-EMPRESA 1, 2, 3

DESCRIPCIÓ-EMPRESA 1, 2, 3

DESCRIPCIÓ-GESTIÓ 1, 2, 3

OPERADOR

OPERADOR-RESUM-GESTIÓ

OPERADOR-VERBAL

BUIT

QUALIFICADOR

A continuació fem una descripció més detallada del tipus ACCIÓ.

En el grup d'accions administratives incluíem tots els verbs que indiquen algun tràmit necessari per una altra acció o que indiquen actes que podríem qualificar d'administratius. Aquests verbs poden presentar com a tema:

- una nominalització d'un verb de tipus acció gestió

- una nominalització d'un verb de tipus acció administrativa
(màxim de recursivitat 2 nivells)
- objectes tals com *ley, convenio, reunión* ...

Sovint el tipus d'agent és un factor discriminador a l'hora d'identificar els verbs d'aquest tipus. El temps verbal més freqüent en aquest grup és el passat (*Pretéritos Perfecto o Indefinido* d'Indicatiu) o el *Futuro Imperfecto* d'Indicatiu.

Són accions discursives aquells verbs propis del tipus del domini que s'utilitzen per introduir informació. Tenen com a tema una nominalització d'un verb de tipus acció de gestió o de tipus acció administrativa. El temps verbal més freqüent és el passat (*Pretéritos Perfecto o Indefinido* d'Indicatiu).

Les accions destructives són, de fet, una particularització de les discursives. Són verbs que desacrediten tot el que ve a continuació doncs les fonts d'informació no són prou fiables. Qualsevol cosa precedida d'un verb d'aquest tipus serà desestimada.

Com accions de gestió entenem tots els verbs més característics del domini. Cada subtipus correspon a una classe de succés econòmic. El temps verbal característic d'aquest grup és el passat (*Pretéritos Perfecto o Indefinido* d'Indicatiu) o el *Futuro Imperfecto* d'Indicatiu.

Qualifiquem com acció de gestió-mínima aquells verbs que descriuen operacions internes dins del marc d'un succés econòmic, o bé que són típics de la descripció dels successos (cas de les comparacions en els resums econòmics). Les associem sempre a un tipus d'acció de gestió. Com són descripcions, el temps verbal per excel·lència és el *Presente* d'Indicatiu.

Finalment, dins del tipus d'acció resultativa queden inclosos els verbs que descriuen accions que es realitzaran després d'un fet econòmic o estats que s'en deriven d'aquest fet. Els trobarem a les notícies qualificades de TEXT-FET, però no a les de TEXT-ESTAT. El temps verbal més freqüent en aquest tipus de verbs és el *Presente* d'Indicatiu o el *Futuro Imperfecto* d'Indicatiu.

A l'Annex es presenta una relació detallada de verbs del domini junt amb el seu tipus. Es pot observar que hi ha verbs que poden pertànyer a diferents tipus segons el tema, l'agent, el temps verbal o l'experimentador. Al mateix temps hi ha verbs que en general poden ser polisèmics, però que en el domini queden totalment definits.

3.3 Reflexions sobre la caracterització realitzada

En aquest punt volem fer algunes consideracions sobre les taxonomies presentades en els apartats anteriors. Creiem que la metodologia aplicada és vàlida tant si es canvia de domini com si s'amplia el domini actual.

El punt de partida ha sigut l'estudi dels textos a tractar, tant en forma com en contingut. Això ha permès establir una tipologia de les notícies a nivell estructural i una altra a nivell conceptual. Al mateix temps ha quedat caracteritzat el coneixement sobre el domini, necessari per la total comprensió del text. Aquest estudi és igualment necessari en qualsevol domini on es vulgui realitzar una comprensió automàtica de textos.

Els textos

La base de coneixement i la tipologia conceptual sempre estaran en funció del domini, com és obvi, però es pot pensar en una tipologia estructural més o menys general, si més no, per les notícies. En cas d'ampliar el domini actual, caldria repasar les tipologies definides, comprovant si apareixen nous tipus conceptuals o bé si l'ampliació comporta alguna estructura de notícia diferent.

La modificació o ampliació de les taxonomies no comporta la modificació del funcionament del sistema. Simplement en el moment de prendre decisions o generar expectatives haurà de tenir en compte l'existència de més alternatives o de noves tipologies i, per tant, haurà de tenir altres criteris de selecció.

Els verbs

La tipologia verbal s'ha fet des d'un punt de vista semàntic i, per tant, depèn del domini. L'ampliació del domini significarà la revisió d'aquesta tipologia: els nous verbs responen a la mateixa classificació, o constitueixen noves classes? El canvi de domini provocarà, amb tota seguretat, una nova tipologia malgrat que es mantinguin les classes més generals. Per exemple, els verbs discursius no són particulars de les notícies sobre economia, els podem trobar a qualsevol tipus de notícies i, fins i tot, a qualsevol text narratiu.

De tota manera, el sistema tractarà els verbs de la mateixa forma en qualsevol domini. Essent que la representació semàntica d'aquests indica si serveixen per crear instàncies o per modificar instàncies (omplir descriptors), només variarà, situats en cada verb, que és el que es crea o el que es modifica.

En resum, el mètode seguit és general encara que, decidit un domini, les propostes de tipologies són particulars.

3.4 Presentació de COTEM: un exemple

Mitjançant l'exposició superficial del processament d'un exemple, donem una visió general del funcionament de COTEM. Utilitzarem la notícia de la figura 3.2, referent a l'acord de creació d'una empresa, com exemple per mostrar els objectius de l'aplicació.

Ens limitem al tractament del titular i el primer paràgraf, doncs és suficient per l'objectiu d'aquest apartat. En el capítol 7 es fa una descripció completa del funcionament de COTEM i es presenten dos exemples detallats de tot el procés.

Reproduïm a continuació la part seleccionada de la notícia.

Corning Glass y Telefónica crean una empresa de fibra óptica

Corning Glass y Telefónica alcanzaron ayer en Nueva York un acuerdo para crear una empresa que fabricará y comercializará fibra óptica en España. Dicho acuerdo está sujeto a la aprobación de los respectivos Gobiernos y consejos de administración.

Considerem l'existència prèvia d'un procés d'anàlisi sintàctica-semàntica que genera una representació del significat del text similar a una gramàtica de casos. Aquest primer procés no ha sigut implementat per no formar part dels objectius del present treball, si bé és factible el seu desenvolupament utilitzant les tècniques descrites a [Rodríguez-89]. L'entrada al sistema serà, doncs, una representació de la notícia mitjançant *frames* verbals.

A partir d'aquesta entrada, el tractament es divideix en dues etapes que, de manera genèrica, anomenem processador i incorporador.

L'objectiu del processador és generar, a partir de la representació inicial del text, els corresponents esquemes segons els conceptes continguts a la Base de Coneixement (conceptes generals, de domini, temporals...). També generarà instàncies que caldrà incorporar al Sistema Temporal. El processador tracta la notícia *frame a frame*. Per realitzar la seva tasca, necessita coneixement sobre els tipus de verbs, els tipus de notícies i la seva combinació. També necessita coneixement lingüístic i de domini per interpretar correctament la informació temporal.

Posteriorment, l'incorporador tractarà d'integrar tots els esquemes generats a la fase anterior, posant en marxa mecanismes de control de compatibilitats i de propagació de relacions temporals.

Representació inicial del text

Els *frames* que representen la part de notícia seleccionada són descrits a continuació de manera simplificada.

F1 correspon al titular.

F1:CREAR [Creació]

AGENTS: EMP1 (*Corning Glass*) i EMP2 (*Telefónica*)

TEMA: EMP3 (*de fibra óptica*)

TEMPS-VERBAL: presente

F2, F3 i F4 representen la segona frase on s'exposa l'acord de creació.

F2:ALCANZAR [Administrativa]

AGENTS: EMP1 (*Corning Glass*) i EMP2 (*Telefónica*)

TEMA: F3

QUAN: "ayer"

LLOC: LOC1 (*Nueva York*)

TEMPS-VERBAL: pretérito indefinido

F3:ACORDAR [Administrativa]

TEMA: F4

F4:CREAR [Creació]

TEMA: EMP3 (*fabricará y comercializará fibra óptica en España*)

F5 i F6 representen la tercera frase, on s'indica que l'acord ha de ser aprovat per altres entitats.

F5:ESTAR-PENDIENTE [Administrativa]

EXPERIMENTADOR: F3

MOTIU: F6

TEMPS-VERBAL: presente

F6:APROBAR [Administrativa]

AGENTS: GOVERN1 (USA) i GOVERN2 (España) i

CONSELL-ADMIN1 (*Corning Glass*) i

CONSELL-ADMIN2 (*Telefónica*)

Processador: Tractament de F1

Considerant el tipus de verb present a F1, el processador dedueix que es tracta d'una notícia de tipus text-fet i que, des del punt de vista conceptual, és una notícia de tipus 3.1.

Del tractament del *frame* com a tal s'obté un nou esquema, instància del concepte "creació", amb alguns descriptors plens:

CREACIO-1

NOVA-EMPRESA: EMP3

PROPIETARIS: EMP1 I EMP2

MOMENT-CREACIO: ET1 (relacionat amb S)

El moment temporal, ET1, és una referència temporal poc precisa, deduïda del temps verbal (fem notar que el present també permet una interpretació com a futur). Aquest ET1 és un element temporal definit per referència a S, la data de la notícia (13 de desembre de 1985).

Processador: Tractament de F2, F3, F4

Del processament conjunt d'aquests tres *frames*, el sistema extreu que el fet de la creació està, ara par ara, en fase d'acord. Per combinació del temps verbal i de l'adverbi temporal, dedueix el moment exacte de l'acord (ET2 = 12 de desembre de 1985).

Per estar en situació d'acord, el sistema pot revisar el moment de creació, ET1. Com no hi ha cap definició concreta sobre el moment de creació, s'usa el coneixement del domini per deduir que ET1 ha de ser posterior a ET2. Aquesta relació és més interessant, de cara a processos de raonament posteriors, que la coneguda fins ara. Per aquest motiu es modifica ET1 deixant-lo relacionat amb ET2 en lloc de la relació anterior amb S.

L'esquema creat abans es veu modificat de la següent manera:

CREACIO-1

NOVA-EMPRESA: EMP3

PROPIETARIS: EMP1 I EMP2

MOMENT-CREACIO: ET1 (relacionat amb ET2)

MOMENT-ACORD: ET2 (valor = 12-12-1989)

LLOC-ACORD: LOC1

L'esquema EMP3 també conté més informació, doncs ara sabem quin és el tipus d'activitat que la nova empresa realitzarà amb la fibra òptica.

Processador: Tractament de F5, F6

Aquests dos *frames* indiquen que l'acord és pendent d'aprobació, així com quines són les entitats que l'han d'aprovar.

Que un acord està pendent d'aprovació, s'interpreta com que alló acordat està pendent d'aprovació. D'altra banda, no tenint més informació temporal, sempre es pot deduir que el moment d'aprovació serà posterior al moment de l'acord. Com hi ha diferents agents per l'aprovació, cal definir un moment temporal per cadascun d'ells, doncs no hi ha cap obligació de que tots decideixin alhora. L'esquema del fet queda ampliat de la següent manera:

CREACIO-1

NOVA-EMPRESA: EMP3

PROPIETARIS: EMP1 i EMP2

MOMENT-CREACIO: ET1 (relacionat amb ET2)

MOMENT-ACORD: ET2 (valor = 12-12-1989)

LLOC-ACORD: LOC1

MOMENT-APROVACIO: (ET3, ET4, ET5 i ET6)

Tots ells posteriors a ET2

SUPERVISORS:

(GOVERN1, GOVERN2, CONSELL-ADMIN1 i CONSELL-ADMIN2)

Incorporador

L'incorporador posa en marxa un conjunt de regles sobre els esquemes obtinguts a la fase anterior per tal de crear i/o modificar objectes de la Base de Coneixement. D'una banda es comprova la coherència de la nova informació amb la continguda en el sistema i, d'altra, tot allò que és nou i coherent és incorporat a la Base de Coneixement i al Sistema Temporal.

L'esquema CREACIO-1, considerat com primera informació al respecte, serà integrat en el conjunt de fets econòmics coneguts. Si EMP1 i EMP2 són empreses ja conegudes, es comprova la compatibilitat de les dades. Tant si es tracta d'empreses conegudes com d'empreses noves que directament cal incorporar, el procediment associat a la creació d'instàncies del fet "creacio" provocarà modificacions sobre EMP1 i EMP2 (descriptor "propietats") i sobre EMP3 (descriptor "creada").

En quant a la informació temporal, l'incorporador fa la unificació d'aquells elements temporals que siguin iguals (només és possible per elements temporals totalment definits) i incorpora tots els elements nous al Sistema Temporal. Durant aquest procés s'activen els mecanismes de propagació de relacions temporals.

Si tot és consistent i la propagació temporal no ha donat errors, queda integrada la notícia.

4. Informació temporal

En el domini de treball el temps és un aspecte fonamental tal com ja s'ha indicat a la introducció. Hi ha moltes dades temporals que cal analitzar per comprendre i incorporar correctament la informació, dades que afecten al raonament posterior. Les notícies es publiquen en una data concreta. Les notícies descriuen fets que s'han produït, es produeixen o es produiran en un moment concret o bé en una certa "distància" temporal respecte un altre fet. La Base de Coneixement s'actualitza a mida que s'introdueixen noves notícies al sistema, per tant va evolucionant en el temps.

4.1 Tipus d'informació temporal en el domini

En el capítol 2 hem parlat dels treballs més importants realitzats fins ara i dedicats a la representació d'aquest tipus d'informació, destacant que gairebé tots tracten tan sols els intervals temporals. Analitzant els textos narratius (les notícies) i, en particular, el domini de treball, és freqüent trobar tant punts com intervals. Per qüestions d'expressivitat opinem que és necessària una representació diferenciada de cada tipus d'element temporal, facilitant-ne així el tractament.

En el nostre domini trobem tres tipus d'informació temporal:

- de caràcter puntual. Per aquesta aplicació considerem que un punt (la unitat mínima de temps) equival a un dia. És evident que hi ha fets econòmics que succeeixen en un dia; per exemple, l'entrada a cotització en Borsa d'una empresa o la decisió d'una ampliació de capital són dades que tenen un dia com element temporal associat.
- de tipus interval. Un interval no és més que una seqüència de dies consecutius i que es defineix indicant el moment (dia, mes. . .) d'inici i el moment (dia, mes. . .) de final o bé fent referència a la totalitat d'un interval predefinit com mes, any o quinzena. Si bé la decisió d'una ampliació de capital es pren en un dia, aquesta operació es porta a terme durant un interval de temps, per exemple, de l'1 al 25 del mes següent al moment de la presa de l'acord.
- de tipus cadena. Una seqüència d'intervals ordenats respecte al temps, però no necessàriament consecutius, constitueix una cadena temporal. Una empresa pot decidir fer una ampliació de capital del 50% de cop, durant un interval de temps, o bé fer la mateixa ampliació en tres etapes, durant tres intervals de temps diferents.

Per ara punts i intervals han sigut implementats. Existeixen pocs treballs sobre cadenes, com ja s'ha exposat en el capítol 2. Deixem com a línia de recerca futura la

implementació de les cadenes que probablement es podrà portar a terme definint un conjunt d'operacions sobre intervals connectats. També es força interessant l'estudi de la seva utilitat per representar les expressions temporals associades a fets iteratius en el temps com per exemple *los lunes*.

4.2 Expressió lingüística del temps

4.2.1 Introducció

Per tal d'obtenir la informació temporal present a una notícia, és necessari un estudi seriós de com s'expressen les nocions temporals en el llenguatge base del text, el castellà en el nostre cas.

Hem tingut en consideració la teoria d'en Reichenbach [Reichenbach-47] que proposa l'existència de tres identificadors, definits a partir del temps verbal, associables a cada frase:

- S (*speech moment*). Moment en el qual s'està produint la frase, per nosaltres és la data d'aparició de la notícia en el diari.
- E (*event moment*). Moment en el qual es produeix el fet que descriu la frase.
- R (*reference point*). Moment que serveix de punt de referència per situar E.

Les úniques relacions que permet establir entre aquests identificadors són: "abans" i "abans o igual". D'aquí s'en dedueix que els moments temporals són considerats de caràcter puntual.

Com ja hem presentat en el capítol 2, aquesta teoria ha sigut utilitzada i ampliada per molts investigadors posteriors. Hornstein fa una proposta de combinació de les estructures E-R-S amb els adverbis temporals. Yip combina el treball d'en Hornstein amb la teoria computacional sobre intervals temporals de l'Allen. En el projecte EUROTRA s'ha passat a considerar intervals en lloc de punts, tractant també l'aspecte verbal.

Fent la consideració d'intervals, les possibles relacions que podem establir són les set definides per James Allen, que es converteixen en tretze si considerem les relacions inverses. Les set relacions bàsiques són: "igual", "abans", "solapa", "durant", "comença", "acaba" i "enllaça".

Com els elements temporals del sistema COTEM poden ser tant punts com intervals, aquestes relacions han sigut especialitzades per tal de distingir les relacions entre intervals, les relacions entre punts i les relacions mixtes.

Quan COTEM interpreta una notícia, recull tota la informació disponible per tal d'assignar a cada fet el seu moment temporal. Per aquest motiu s'ha estudiat tant els temps verbals com qualsevol altre tipus d'expressió de la informació temporal.

COTEM incorporara mecanismes de combinació del E-R-S dels temps verbals amb els adverbis temporals i les locucions i expressions que també representen directament o indirectament el temps.

El mode, temps, aspecte i altres característiques del verb ens permeten saber si una acció és present, futura o passada, si és tan sols una possibilitat, si està acabada o no... El verb serveix per definir l'ordre entre E, R i S. Els adverbis temporals (avui, demà, sempre) ajuden a concretar el temps de l'acció, modificant a vegades el sentit temporal del verb (Demà obrim a les nou). Igual que els adverbis, altres expressions afegixen una noció temporal a la frase (el 23 d'abril, la setmana passada, durant un mes, després de la compra d'accions).

La implementació actual s'ha definit tenint en compte les següents restriccions lingüístiques:

- L'aspecte verbal només ha sigut considerat per alguns verbs que són qualificats com operadors sobre el següent verb o nominalització. És el cas del verb *prever* o del verb *realitzar*. Per la resta de casos només és té en consideració el temps verbal.
- En quant a la combinació entre els temps verbals i les expressions temporals, només es consideren els casos on l'expressió no canvia l'ordre de la terna E-R-S definit pel temps verbal, és a dir, els casos on la combinació serveix per concretar més les relacions E-R-S i per donar valor a E o R (*La empresa X ampliarà su capital social durante el próximo marzo*).

D'altra banda existeixen també restriccions pròpies del tipus d'aplicació:

- Per COTEM l'identificador S és sempre un punt (la data de la notícia). El tipus d'E depèn del fet que s'està descrivint mentre que R acostuma a ser un interval.
- Les expressions temporals ens defineixen punts (avui, el 23 d'abril), intervals (des de l'1 de gener al 15 de febrer, durant la setmana passada, des de fa tres anys, el proper mes) o referents a altres elements del sistema temporal (després de la compra d'accions). Les expressions temporals corresponen a E o a R. En el cas de que l'expressió sigui un referent a un altre element temporal, pot ser que la referència es pugui resoldre en el propi text o que sigui necessària una consulta a la Base de Coneixement. Aquest tipus d'expressions no han sigut tractats de moment doncs requereixen tècniques de resolució de referents en el text que en cas de fracàs han d'activar la cerca de referents dins la Base de Coneixement. El desenvolupament d'aquestes tècniques queda fora dels objectius de recerca actuals.

- Un problema addicional és el tractament d'expressions que fan referència a un període de temps indeterminat, per exemple: alguns anys, uns quants mesos, primers de mes... Algunes d'elles han sigut implementades com és el cas de primers, mitjans o finals de mes/any.

El domini també pot aportar restriccions. En el cas de la informació temporal s'ha observat que hi ha expressions temporals particulars d'algun tipus de notícia, com es pot veure a l'apartat de descripció detallada d'aquestes expressions.

A continuació, subapartat 4.2.2, es descriu l'estudi realitzat sobre els temps verbals del castellà i l'assignació de l'estructura E-R-S a cadascun d'ells que utilitza COTEM. En el subapartat 4.2.3 presentem les expressions temporals i la tipologia definida. I en el subapartat 4.2.4 parlem de la combinació dels tipus de verbs, els temps verbals i les expressions temporals. Es mostren uns exemples de combinació, però el detall d'aquest procés s'exposa en el capítol 7. El subapartat 4.2.5 presenta algunes conclusions sobre el treball lingüístic desenvolupat.

4.2.2 Assignació de l'estructura ternària (E-R-S) als temps verbals

Aquesta definició s'ha realitzat a partir de la teoria sobre els tres elements temporals desenvolupada per Reichenbach, presentada a l'apartat anterior. Com ja hem comentat, COTEM treballa amb intervals i punts motiu pel qual les relacions temporals possibles entre els elements d'una estructura ternària són les tretze definides per Allen, però especialitzades segons els tipus dels elements relacionats.

La nostra proposta [CaMaVe-89a] està en línia amb els treballs realitzats pel grup bidisciplinar de Toulouse [BoBoBras-88] si bé ells consideren només que E, R i S són intervals.

Els patrons verbals es defineixen de manera genèrica, és a dir, definim les relacions entre els tres identificadors com si fossin intervals. En el moment de la representació d'elements temporals concrets en el Sistema Temporal es fa la distinció entre punts i intervals i per tant l'especialització de les relacions.

Distingir entre punts i intervals permet freqüentment reduir el nombre de relacions presents en el patró. Per exemple, si el patró diu que la relació entre E i R ha de ser una relació múltiple tal que E està contingut en R o bé E és igual a R i, a més, del text i/o el domini s'ha deduït que tant E com R són punts, la relació múltiple queda convertida en una única relació que és la igualtat entre punts. Per tant el nombre de relacions emmagatzemades es redueix, augmentant així la precisió de la informació que conté el sistema. Encara més, la distinció entre tipus d'elements millora l'eficiència del sistema doncs resulta que la reducció de relacions mantingudes entre els elements temporals provoca la reducció del temps d'execució de l'algorisme de

propagació de relacions. Aquest algorisme, a l'igual que tot el Sistema Temporal, es descriu amb detall en el capítol 6.

Un sistema de comprensió de notícies ha de contenir una representació dels temps verbals. Ara bé, segons el domini seleccionat, pot decidir que alguns temps verbals no cal representar-los doncs no surten mai. A més la representació d'un temps verbal pot variar segons el domini. No obstant, els mecanismes de tractament d'aquestes representacions es poden definir de manera genèrica.

En el nostre domini (notícies d'economia) no apareixen tots els temps verbals. Hi ha alguns temps molt utilitzats (Indicatiu: *Presente*, *Pretérito Perfecto*, *Pretérito Indefinido*, *Futuro Imperfecto*), altres que apareixen en una certa freqüència (Indicatiu: *Pretérito Imperfecto*, *Futuro Perfecto*), altres poc freqüents (*Pretérito Anterior* d'Indicatiu, *Presente* de Subjunctiu, *Condicional*) i altres, els no esmentats, que no surten mai. Fins i tot, segons el tipus de notícia, podem definir un temps verbal predominant (*Futuro Imperfecto* d'Indicatiu per les notícies sobre creacions d'empreses o *Pretérito Perfecto* d'Indicatiu a les notícies sobre compres). També podem definir temps verbals associats a les parts de les notícies; per exemple, *Presente* o *Pretérito Imperfecto* d'Indicatiu per la part descriptiva i *Pretérito Perfecto* o *Pretérito Indefinido* o *Futuro Imperfecto* d'Indicatiu per la part d'informació principal. Ateses aquestes característiques ens hem limitat als temps més freqüents del mode Indicatiu.

Abans de passar a l'exposició de les estructures ternàries per cada temps verbal, cal fer dues observacions de caràcter general.

- 1) Recordem que el sistema té com unitat temporal mínima el dia. Aixó fa que algunes relacions entre els tres identificadors, que semblarien possibles des d'un punt de vista lingüístic ampli, no són considerades per no ser factibles quan el dia és el nivell mínim. Per aquest motiu també apareixen les relacions d'enllaça i enllaçat-per que en altres treballs no hi consten.
- 2) Cal tenir en compte que COTEM defineix S com la data de la notícia mentre que les notícies són redactades el dia abans tenint en consideració que la gent les llegirà al dia següent de la seva redacció. Aixó provoca que s'hagi d'admetre alguna relació de més en els patrons que pot semblar artificiosa, però que és necessària per tal de que la combinació entre temps verbals i expressions temporals funcioni.

Codificació de les relacions (a Rel b) entre intervals:

<	abans	>	després
e	enllaça	ei	enllaçat per
s	solapa	si	solapat per
d	durant	di	conté
c	comença	ci	començat per
a	acaba	ai	acabat per
=	igual		

En el cas de punts, la codificació es redueix a dues possibilitats:

- abans , abans o igual

A continuació presentem els patrons verbals utilitzats pel sistema, comparant-los amb les propostes d'altres investigadors.

Representació del Presente d'Indicatiu:

Hornstein proposa: S,R,E

Yip admet S,R,E i E,R,S per tal de resoldre alguns problemes plantejats per Hornstein.

Sembla que la situació E,R,S té sentit en castellà si baixem al nivell temporal de les hores, aquest no és el cas en el nostre sistema (unitat temporal mínima = dia). Per tant la nostra proposta és:

S Rel E	e < = d c a
S Rel R	e < = d c a
E Rel R	= d c a

Aquesta proposta coincideix amb la de [BoBoBras-88].

Representació del Pretérito Perfecto d'Indicatiu:

Hornstein proposa: E,S,R

Yip proposa: E,R,S

Atenent a les característiques del castellà, ens sembla més adequada la proposta d'en Yip. Tenim dues representacions possibles segons si a la frase ens defineixen R mitjançant alguna expressió temporal o no. Si no hi ha definició de R, el referent és un interval des de qualsevol moment fins just S.

S Rel E	> ei
S Rel R	ei
E Rel R	= d c a

Si hi ha definició de R, sempre serà el cas de que R conté S d'alguna manera i també R conté E d'alguna manera. Més concretament:

$$\begin{array}{l} S \text{ Rel } E > ei \\ S \text{ Rel } R d a \\ E \text{ Rel } R d c \end{array}$$

Aquesta segona opció és gairebé igual a una (*passé composé* 1) de les proposades a [BoBoBras-88], només es diferencia en que a la relació entre S i E nosaltres hem afegit ei.

De cara a definir un patró, és millor fer una representació única que durant el procés d'adaptació ja quedarà reduïda de manera escaient. El patró utilitzat és:

$$\begin{array}{l} S \text{ Rel } E > ei \\ S \text{ Rel } R d a ei \\ E \text{ Rel } R = d c a \end{array}$$

Representació del Pretérito Imperfecto d'Indicatiu:

No tenim proposta de Hornstein ni de Yip atés que l'anglès no té aquest temps. Tampoc tenim la versió de A. i M. Borillo i M. Bras doncs sembla que aquest temps no l'han treballat.

La nostra proposta és:

$$\begin{array}{l} S \text{ Rel } E > ei \\ S \text{ Rel } R > ei \\ E \text{ Rel } R = \end{array}$$

Aquest patró és així atés el domini, doncs l'imperfet s'utilitza per descripcions d'estats. Aquest temps també s'usa per fets iteratius mirats com un tot, però no s'ens presenta aquesta opció en el domini. El més freqüent és que E sigui un interval.

Representació del Pretérito Indefinido d'Indicatiu:

Hornstein i Yip proposen: E,R_S

Si el verb està en Indefinit, sempre surt a la frase una expressió temporal que serà un posicionament o una definició del R.

$$\begin{array}{l} S \text{ Rel } E > ei \\ S \text{ Rel } R > ei \\ E \text{ Rel } R = d c a \end{array}$$

Aquesta proposta és semblant a la poposada a [BoBoBras-88] com segona opció del *passé composé*, si bé hem afegit la relació ei entre S i E i entre S i R.

Representació del Futuro Imperfecto d'Indicatiu:

Hornstein i Yip proposen: S_R,E

No tenim la versió de A. i M. Borillo i M. Bras.

La nostra proposta és:

S Rel E	< e
S Rel R	< d e c
E Rel R	= d c a

Representació del Futuro Perfecto d'Indicatiu:

Hornstein i Yip proposen: S_E_R

D'aquí s'en deduiria:

S Rel E	<
S Rel R	<
E Rel R	>

Aquesta proposta coincideix amb [BoBoBras-88].

Nosaltres hem d'afegir l'opció d'enllaça doncs la unitat temporal mínima és el dia. Per tant:

S Rel E	< e
S Rel R	< e
E Rel R	> ei

4.2.3 *Les expressions temporals: Tipologies*

S'ha fet un estudi exhaustiu de la mostra per tal de tipificar totes les expressions temporals que hi apareixen. D'acord a la seva semàntica i al procés d'interpretació que cal realitzar per cadascuna d'elles, podem establir la següent classificació:

- expressions temporals que es poden interpretar per sí mateixes o bé amb l'ajut d'alguna altra expressió temporal. Són les que hem classificat i tractat.
- expressions temporals que no afegixen més informació que la que podem obtenir del propi temps verbal, com *ahora*, *en la actualidad*, *próximamente*, etc. En aquest cas l'analitzador s'encarrega d'eliminar-les doncs són irrelevantes.
- expressions temporals que indiquen iteració d'un fet, per exemple *ventas anuales*, *reunión mensual*, etc. La implementació actual de COTEM no pot interpretar-

les, però els elements de tipus cadena poden ser una bona solució per aquestes expressions.

- expressions temporals relatives, és a dir, aquelles que presenten una referència temporal a un altre fet diferent d'aquell del qual s'ens està informant. Per resoldre-les cal cercar en el text de la notícia i/o a la Base de Coneixement. En una nova versió de COTEM, un cop ja experimentat el tractament dels elements temporals en general, afegirem aquesta opció.
- altres expressions, no tipificades, de l'estil de *para los futuros tiempos de la libre competencia, en sus dos años de actividades, tras seis ejercicios con pérdidas...* De moment no s'analitzen.

De les 409 expressions temporals presents a la mostra, 353 corresponen al primer grup, 10 corresponen al segon grup, 3 corresponen al tercer grup, 17 corresponen al quart grup i 26 corresponen al cinqué. Tenint en compte que els tres darrers grups no tenen tractament encara, COTEM interpreta actualment el 88,75% de la mostra pel que fa a aquest tipus d'expressions.

Presentem a continuació l'estudi detallat de les expressions temporals que corresponen al primer grup.

Des del punt de vista de la sintàxi, existeixen dos tipus d'expressions temporals: les que acompanyen al verb i les que acompanyen a un grup nominal. Les distingirem en el moment de la representació sintàctica-semàntica associant les primeres al cas QUAN del *frame* verbal corresponent mentre que les segones seran associades al cas del tema, de l'agent o de l'experimentador segons correspongui. El procés de càlcul que converteix les expressions a elements temporals és el mateix per tots dos tipus. En canvi, el procés d'assignació d'aquests elements temporals a objectes de la Base de Coneixement és diferent: el criteri per decidir si es tracta d'una definició per E o per R es fa en base al tipus d'element temporal obtingut i no pas en base al tipus de quan, tal com s'exposa en el subapartat 4.2.4.

Les expressions temporals tractades s'interpreten com a punts, intervals definits o vaguetats. Aquestes últimes han sigut tractades parcialment com ja hem comentat (cas de *primeros de año*), si bé en els arbres d'anàlisi estan totes recollides.

A continuació presentem la classificació d'expressions temporals utilitzada per COTEM. Aquesta classificació s'ha fet en base a l'estructura sintàctica de les expressions temporals, afegint algunes restriccions semàntiques com és el fet de considerar si l'expressió defineix un punt o un interval. Per cada tipus fem una descripció de les seves característiques i mostrem el seu arbre d'anàlisi on es pot observar fàcilment quines són les expressions analitzades. Una primera aproximació fou presentada a [Castell-88b]. En

el report [CaMaVe-89b] es pot consultar una descripció més detallada de tot el treball lingüístic realitzat.

Per tal de facilitar la comprensió dels arbres d'anàlisi farem algunes explicacions sobre el format:

- Els arbres consten de nodes terminals, identificats per $\langle *nom-node \rangle$, i nodes no terminals, identificats per $\langle nom-node \rangle$.
- Els nodes terminals són els que ja corresponen directament al text analitzat i a les figures apareixen alguns exemples de valors possibles per cadascun d'ells. El node terminal $\langle *NOMT \rangle$ s'utilitza de manera general encara que dels exemples es pot deduir que no sempre admet el seu sentit més ampli. Una situació similar es produeix amb $\langle *NUM \rangle$.
- Els nodes no terminals són nodes intermitjos que encara presenten alguna subestructura sintàctica.
- Tant els nodes terminals com els no terminals poden ser opcionals. En aquest cas en lloc d'indicar el nom entre $\langle \rangle$ apareix entre $[]$.
- Alguns nodes terminals dels arbres tenen com a valor directament un literal que acostuma a ser una preposició. Una expressió temporal serà acceptada per un arbre d'anàlisi que contigui literals si posseeix el mot o mots indicats i en el lloc indicat, apart de complir totes les restriccions dels altres nodes.
- Situats en un nivell qualsevol de l'arbre, l'ordre d'esquerra a dreta és relevant. Els adjectius són una excepció d'aquesta regla. En el text podem trobar *el mes pasado* i *el pasado mes*, mentre que en els arbres hem posat sempre el node que designa els adjectius a l'esquerra del nom corresponent. Per tant totes dues expressions queden recollides sota el mateix arbre d'anàlisi.
- Si un arbre aprofita la definició completa d'un altre, s'indica posant el nom d'aquest segon arbre dins d'un rectangle.
- Quan en una expressió apareix un mot que correspon a un període de calendari seguit del seu valor, tots dos elements són agrupats formant un únic valor. Per exemple, l'expressió *el próximo mes de enero* es tracta com si fos *el próximo mes-enero* i l'expressió *el ejercicio de 1989* es tracta com *el ejercicio-1989*.

Passem a la descripció individualitzada de cada tipus d'expressió temporal.

QUAN-0

En aquest quan queden representades les expressions temporals sense preposició que especifiquen un element de calendari, és a dir, que directament o amb algun càlcul senzill podem deduir-ne un dia, un mes, un any, un semestre... (totalment definits). El QUAN-0 també representa setmanes i dies de la setmana. Els dies de la setmana

són traduïts al corresponent dia del mes. Per les setmanes es calculen els extrems corresponents, a nivell de dies, i es genera un interval. Aquí es considera el mot *ejercicio* sinònim d'any natural.

Observant l'arbre d'anàlisi es constata que les expressions temporals d'aquest tipus consten d'una part obligatòria, la unitat temporal explícita, i una part opcional, el marc temporal. La part obligatòria conté uns modificadors opcionals (determinants, adjectius temporals i adjectius ordinals) i un nucli temporal obligatori que pot consistir en la referència a un dia, un mes, un trimestre, etc. Aquest nucli temporal pot estar situat en relació a un altre període temporal que l'engloba, el marc temporal. Aquesta part opcional s'introdueix per la preposició *de* i ja presenta algunes restriccions: si el nucli temporal fa referència a un any, no pot haver-hi marc temporal; si el nucli temporal fa referència a un dia, el marc temporal ha de ser una setmana o un mes;... Només s'admeten les combinacions raonables. Si dins del primer nivell del marc temporal no s'està fent referència a un any, l'expressió temporal encara pot tenir una altra preposició *de* per introduir la descripció de l'any.

Combinant els valors possibles indicats en els nodes terminals, es poden formar molts exemples d'aquest tipus d'expressió: *el próximo dia 5, este mes, el primer trimestre de 1988, el 17 de junio del pasado año...*

El QUAN-0 serveix de base per la definició de molts altres tipus d'expressions com podem observar en la resta dels arbres d'anàlisi. En alguns d'aquest arbres trobarem els nodes QUAN-0-P i QUAN-0-I que indiquen la presència d'un QUAN-0, però restringit a que sigui un punt o un interval, respectivament.

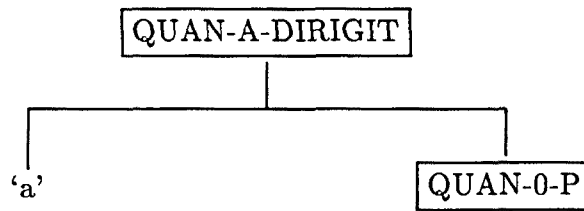
Va associat al verb.

QUAN-A-DIRIGIT

Aquest quan correspon a les expressions temporals que comencen per la preposició *a* i contenen una expressió específica que defineix un punt (un QUAN-0 restringit a punt). Té la particularitat de que serà convertit a un interval en funció de l'element econòmic del qual s'estigui parlant.

Exemples: *a 31 de diciembre del pasado año, a 1 de octubre...*

Va associat al verb.

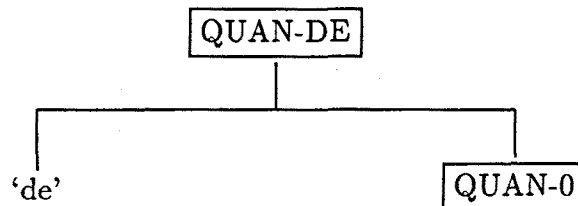


QUAN-DE

En aquest quan queden representades totes les expressions temporals que comencen per la preposició *de* i que especifiquen el mateix que en un QUAN-0.

Exemples: *del pasado enero, de este jueves, de 1987...*

Va associat a un grup nominal.

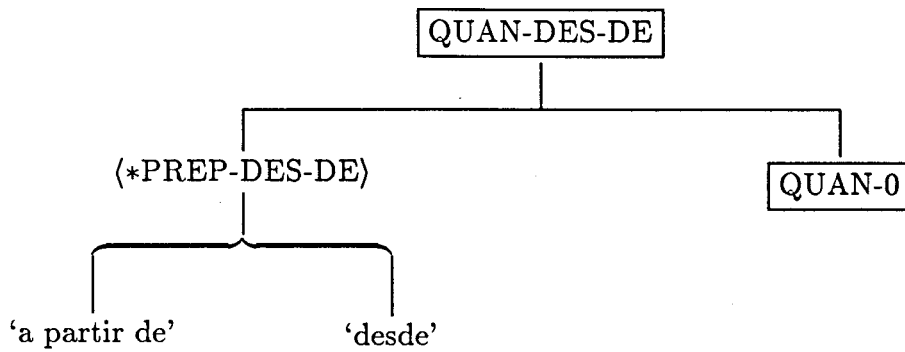


QUAN-DES-DE

En aquest quan queden representades totes les expressions que comencen per *a partir de* o *desde* i contenen una expressió específica (un QUAN-0). Totes les expressions d'aquest tipus trobades a la mostra contenen una expressió que defineix un punt, però no hem restringit el QUAN-0 per tal de mantenir la simetria entre aquest quan i el QUAN-FINS. Defineixen un interval del qual només sabem quan comença.

Exemples: *desde hoy, a partir del 15 de setiembre, a partir de este año...*

Va associat al verb.

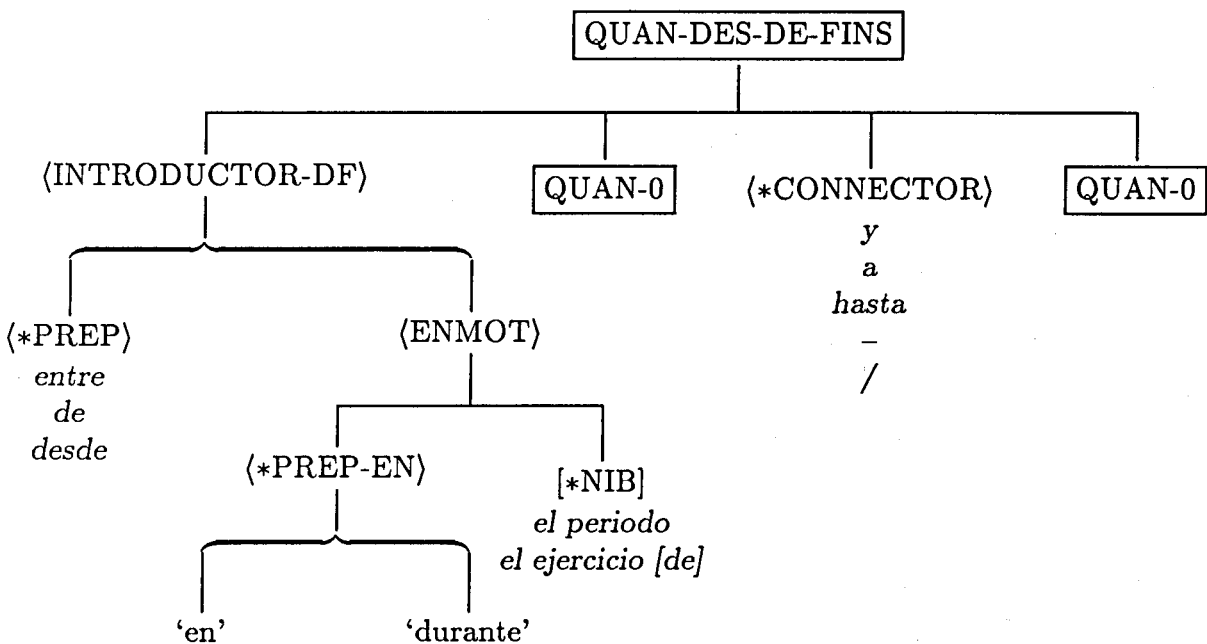


QUAN-DES-DE-FINS

Aquest quan agrupa les expressions en les quals es defineix un interval donant els dos extrems. Els extrems poden ser a la vegada altres intervals, no es obligatori que siguin punts. Té la particularitat de representar anys econòmics que cal interpretar com intervals d'octubre a setembre i no com dos anys.

Exemples: *de enero a octubre, en el ejercicio 1987/1988, desde el 2 de setiembre al 5 de octubre...*

Va associat al verb.

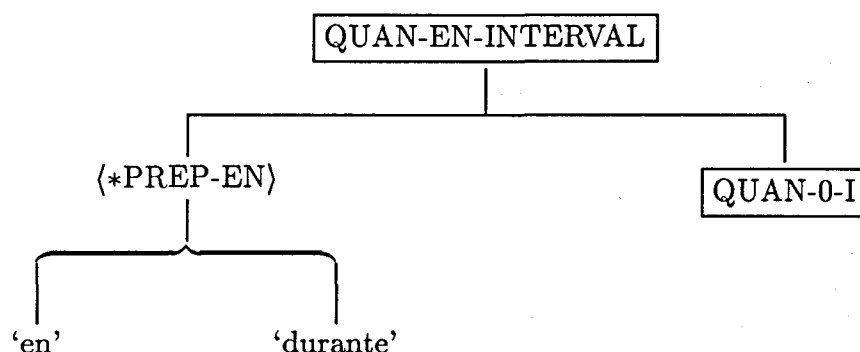


QUAN-EN-INTERVAL

Aquest *quan* correspon a les expressions temporals que comencen per preposicions com *en* i *durante* i contenen una expressió específica que denota un interval (un QUAN-0 restringit a interval). És, amb molta diferència, l'expressió temporal més utilitzada en el domini.

Exemples: *en el pasado año, durante el mes de mayo, en el primer semestre...*

Va associat al verb.



QUAN-EN-LAPSE

S'inclouen en aquest tipus les expressions temporals que defineixen un interval de temps en funció d'unitats temporals. Simplificant, podem tenir un número i una unitat temporal (*años, meses...*) o un lapse no ben definit. El segon cas s'hauria pogut tractar amb valors per defecte, però no ha semblat massa raonable.

Exemples: *en los próximos tres meses, en el plazo de 10 días, en los últimos años...*

Va associat al verb.

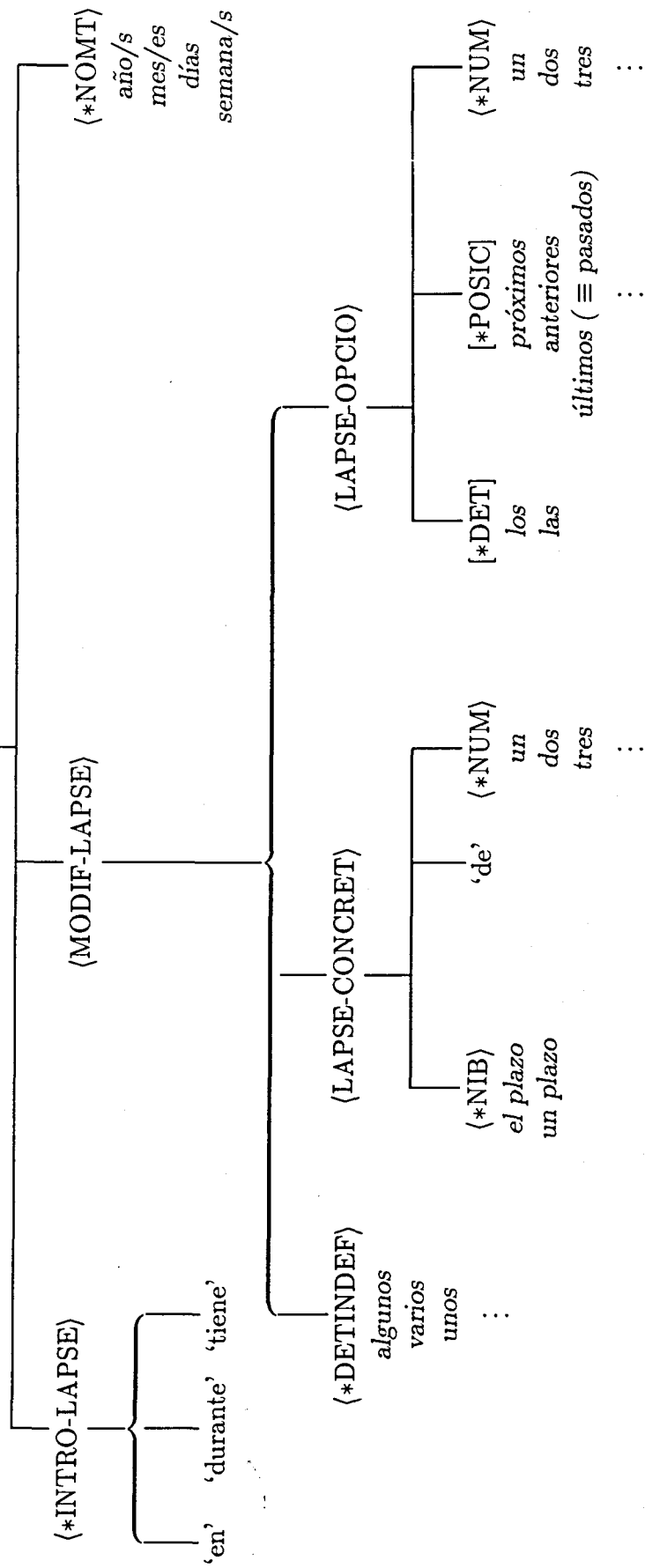
QUAN-FA

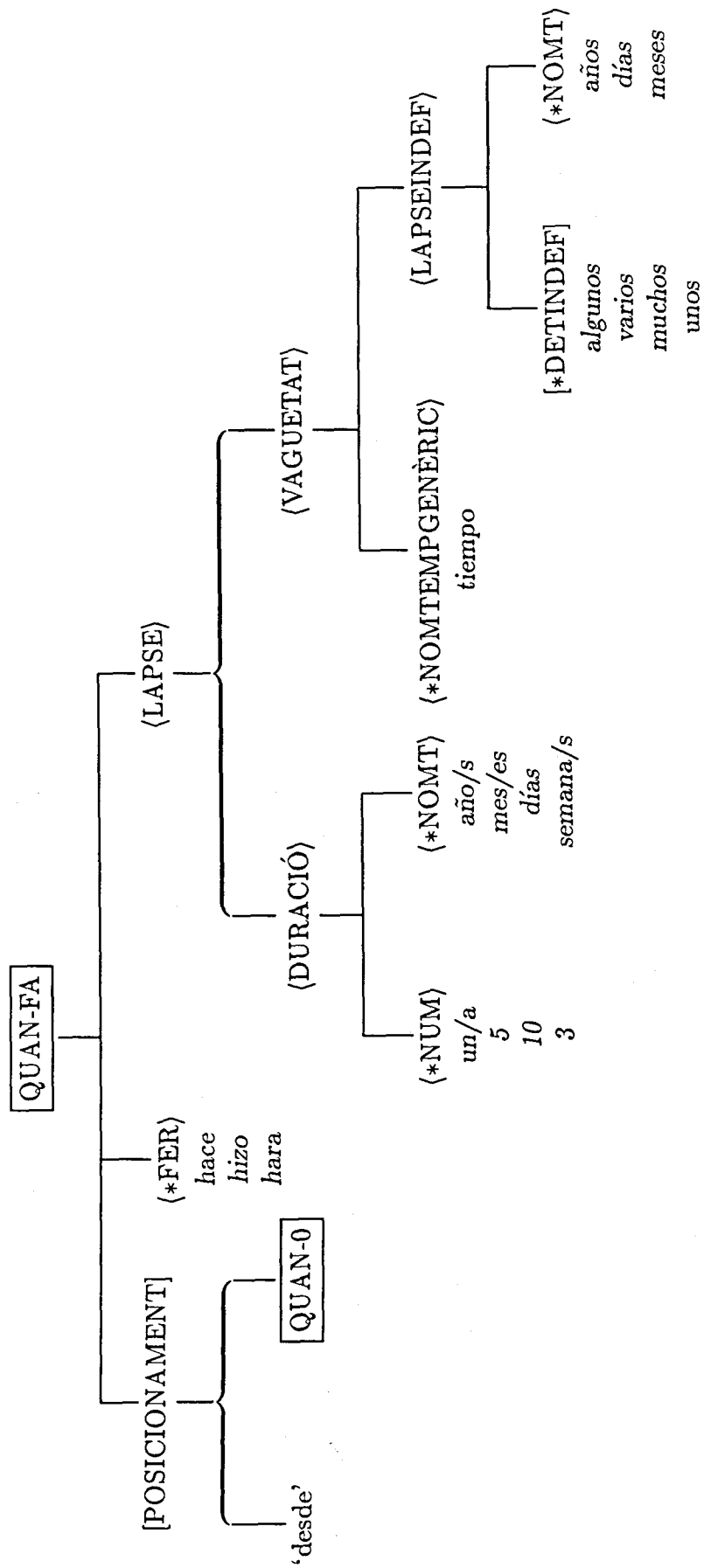
Amb aquest *quan* analitzem totes les expressions que contenen el verb *fer* en les tres opcions impersonals: *hace, hizo* i *hará*. Com es pot veure a l'arbre d'anàlisi, presenta una opció de vaguetat que hauriem pogut implementar prenent valors per defecte, però no ha semblat massa raonable. El posicionament és obligatori si la forma impersonal no és *hace*. Del lapse que defineixen aquestes expressions ens pot interessar només l'extrem o tot el lapse. El criteri seguit, segons la mostra, és quedar-se amb l'interval quan l'expressió comença amb *desde* i prendre només l'extrem en els altres casos.

Exemples: *hoy hace tres meses, desde hace años, el lunes hará cinco años...*

Va associat al verb.

QUAN-EN-LAPSE



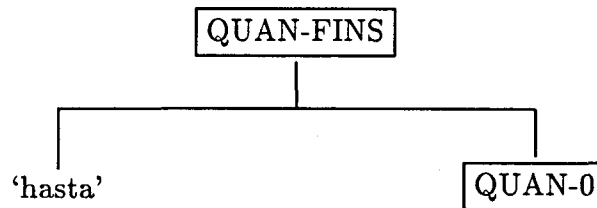


QUAN-FINS

En aquest *quan* queden representades totes les expressions que comencen per *hasta* i contenen una expressió específica (un QUAN-0). Defineixen un interval del qual només sabem quan acaba.

Exemples: *hasta el próximo julio, hasta ayer, hasta el año 1988...*

Va associat al verb.

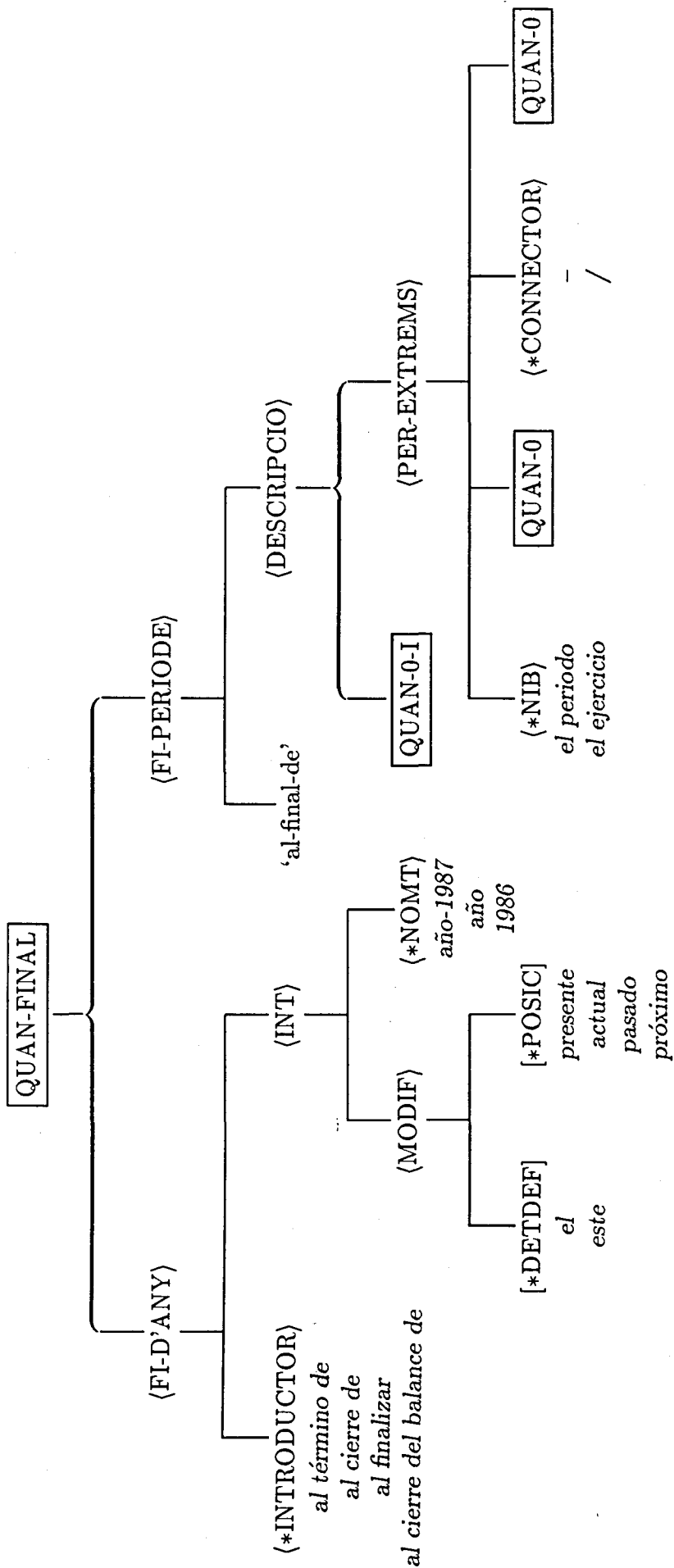


QUAN-FINAL

Aquest *quan* correspon a les expressions que ens situen al final d'un any o d'un període que pot estar definit per un QUAN-0 (restringit a interval) o per dos extrems. Encara que ens situï al final d'un interval, pot ser que ens interessen tot l'interval segons l'element econòmic del qual s'estigui parlant. Té la particularitat de representar períodes que són anys econòmics com intervals des d'octubre a setembre i no com dos anys.

Exemples: *al término del presente año, al final del periodo enero-octubre, al final del primer trimestre...*

Va associat al verb.

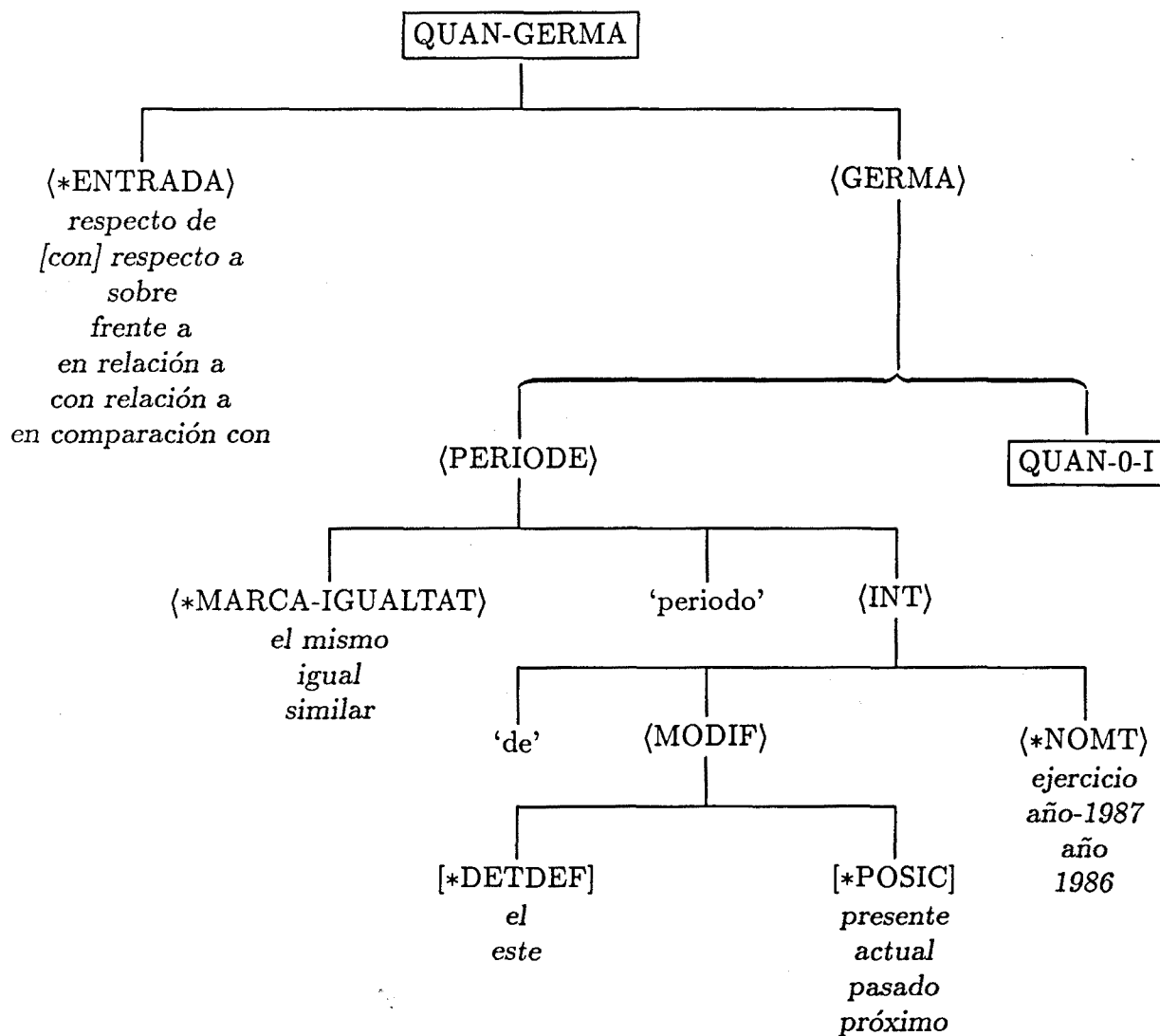


QUAN-GERMÀ

Aquest és un quan molt particular del domini i, encara més, d'un tipus de notícia concret (les que parlen de resums econòmics). Són expressions que comencen amb *respecto de*, *en relación a*, o similars, seguit d'un interval ben definit (un QUAN-0 restringit a interval) o d'una referència a un altre període (el germà). En tots dos casos es necessita l'element temporal germà per tal de calcular correctament la representació de l'expressió *el año anterior*.

Exemples: *con respecto al mismo periodo del año anterior*, *en relación al pasado año...*

Va associat al verb.



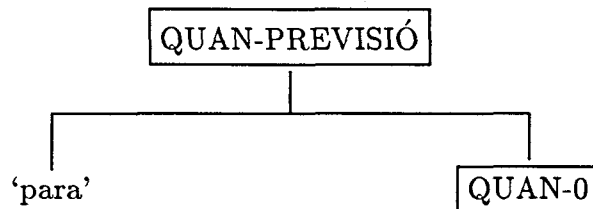
QUAN-PREVISIÓ

En aquest quan queden representades totes les expressions que comencen per la

preposició *para*. És un equivalent al QUAN-DE que denota previsió.

Exemples: *para el próximo lunes, para este año, para el primer semestre...*

Va associat a un grup nominal.

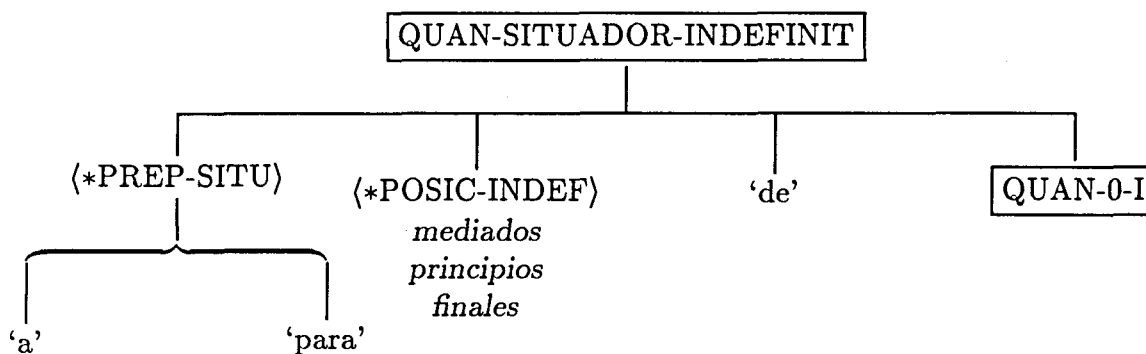


QUAN-SITUADOR-INDEFINIT

En aquest *quan* queden representades les úniques expressions de vaguetat que en aquest moment tracta el sistema. Són expressions que comencen per *a* o *para* (previsió) i es refereixen a *primeros, mediados* o *finales* d'un interval ben definit (un QUAN-0 restringit a interval), que normalment és un mes o un any.

Exemples: *a mediados de año, a principios de mes, para finales de año...*

Va associat al verb.

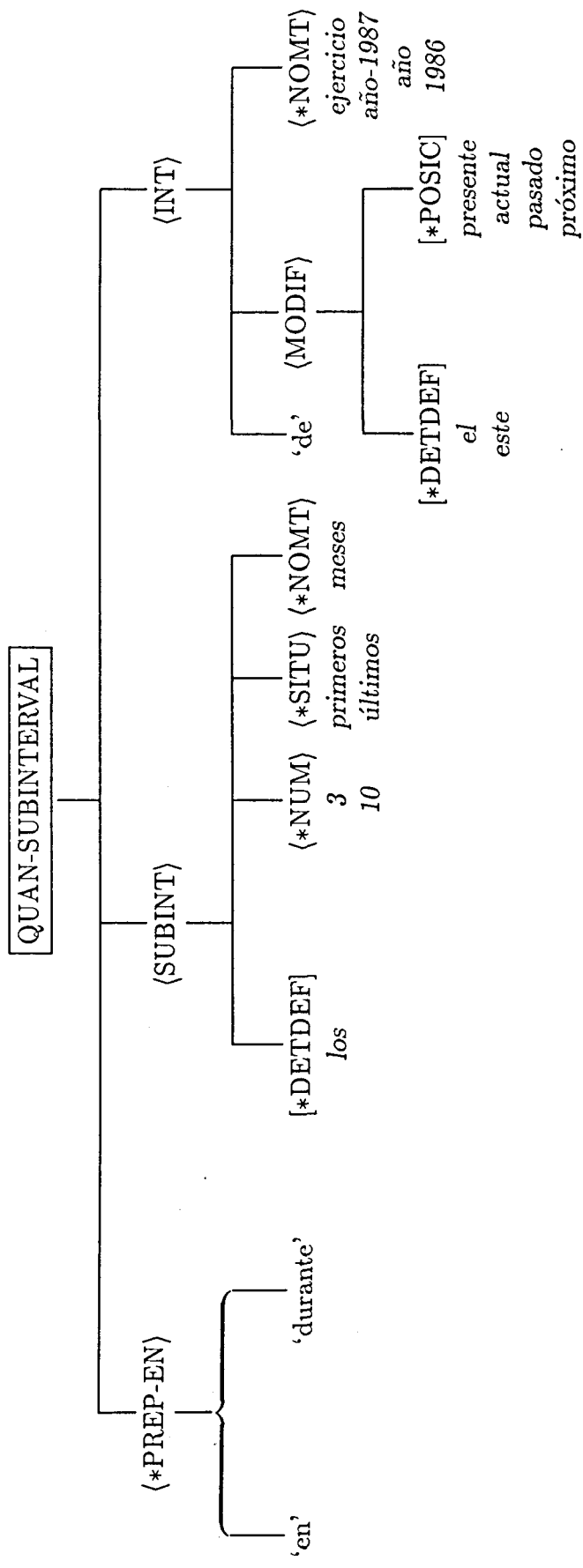


QUAN-SUBINTERVAL

En aquest *quan* queden analitzades les expressions temporals que defineixen un lapse de temps com a subinterval d'un element de calendari, més concretament defineixen un lapse exacte de mesos respecte d'un any (que si és possible es representa com trimestre o semestre). Aquí es considera el mot *ejercicio* sinònim d'any natural.

Exemples: *durante los cinco primeros meses, los dos últimos meses de 1987...*

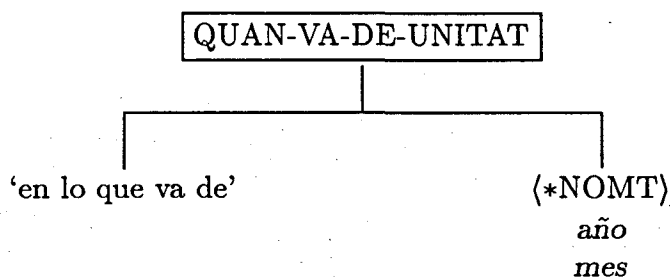
Va associat al verb.



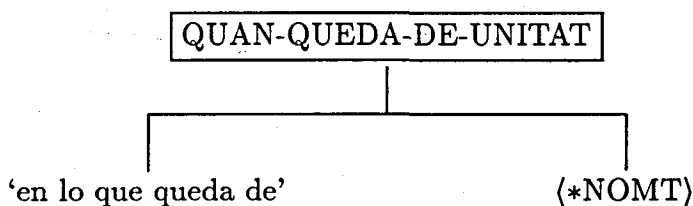
QUAN-VA-DE-UNITAT

Aquest quan correspon a un tipus d'expressions temporals molt peculiars. Són expressions que comencen per *en lo que va de* i a continuació hi ha una unitat temporal (restringida de fet a *mes* o *año*). Defineixen doncs un interval des de l'inici de l'unitat fins al moment actual.

Va associat al verb.



El cas complementari és el QUAN-QUEDA-DE-UNITAT per les expressions *en lo que queda de* + unitat temporal. Aquest cas no apareix a la mostra.



4.2.4. Relació entre expressions temporals, temps verbals i tipus de verbs.

Feta la classificació de les expressions temporals, hem cercat possibles relacions entre aquesta taxonomia i els diferents temps verbals. No s'ha trobat cap relació directa entre algun tipus de *quan* i un temps verbal particular.

En canvi, l'estudi de possibles relacions entre tipus d'expressions temporals i tipus de verbs sí ha donat alguns resultats interessants. Aquests resultats ens permeten simplificar el procés de combinació del temps verbal amb l'expressió temporal de cara a obtenir la interpretació temporal global d'una frase. Recordem que el QUAN-DE i el QUAN-PREVISIÓ no van associats al verb sinó a un grup nominal, segons el tipus d'element temporal que defineixen podem dir que directament corresponen al fet (E) o són un referent (R).

Hi ha un grup d'expressions temporals que són de caràcter general i poden aparèixer acompanyant a qualsevol tipus de verb. Són les expressions que corresponen a les classes QUAN-0, QUAN-EN-LAPSE, QUAN-FA i QUAN-SITUADOR-INDEFINIT.

Considerem que aquestes expressions defineixen R i que les podem trobar a qualsevol domini. Pot ser que en un altre domini tinguin altres interpretacions.

Un segon grup el constituïxen expressions temporals que, sense perdre un cert caire general, són més pròpies d'algun tipus de verb. Són les expressions de les classes QUAN-DES-DE-FINS (sobretot per verbs d'ampliació-de-capital i de resum-gestió), QUAN-FINS (sobretot per verbs d'ampliació-de-capital i de resum-gestió) i QUAN-EN-INTERVAL (sobretot per verbs de resum-gestió). Quan aquestes expressions apareixen acompanyant algun dels verbs indicats entre parèntesis, defineixen E i, en la resta de casos, defineixen R. També aquest grup d'expressions és vàlid per altres dominis, si bé igualment pot succeir que les interpretacions siguin diferents.

Finalment, hi ha un tercer grup d'expressions temporals que sempre apareixen acompanyant a un tipus de verb en particular. Les expressions de la classe QUAN-DES-DE sempre acompanyen verbs d'ampliació-de-capital. Les expressions de les classes QUAN-A-DIRIGIT, QUAN-FINAL, QUAN-GERMÀ, QUAN-SUBINTERVAL i QUAN-VA-DE-UNITAT són típiques dels verbs de resum-gestió. Aquestes expressions defineixen E. Tenim un grup d'expressions, totalment típiques del domini, que en altres dominis tindran gairebé segur una interpretació diferent i, fins i tot, pot ser que no es presentin en altres dominis (per exemple el QUAN-GERMÀ és molt i molt peculiar del domini).

El mecanisme de combinació de la informació temporal es presenta de manera detallada en el capítol 6. A continuació, només a títol il·lustratiu, presentem uns exemples d'aquest procés de combinació.

Exemple-1:

Manfred Heckle, director del área de ..., anunció en Francfort que su empresa y Siemens fundarán en 1987 una firma conjunta que se dedicará a la producción ...

Per definir el moment temporal de la creació de la nova empresa utilitzarem el verb (*fundarán*) y l'expressió temporal (*en 1987*).

Pel temps verbal sabem que

S Rel E < e

S Rel R < d e c

E Rel R = d c a

Per la notícia sabem que S és el 7 de novembre de 1986.

L'expressió temporal és de la classe QUAN-EN-INTERVAL i dels comentaris anteriors es dedueix que defineix R.

Per la Base de Coneixement sabem que E ha de ser un punt (moment de creació).
De la combinació de totes aquestes dades resulta:

$$\begin{aligned} S \text{ Rel } E &< \\ S \text{ Rel } R &< \\ E \text{ Rel } R &= d \text{ c a} \end{aligned}$$

i la informació que el sistema ha de guardar per tenir ben definit (amb el que coneix fins ara) el moment de creació és R (en 1987), E (punt) i la relació E-R durant, comença o acaba (convenientment adaptada al cas punt-interval).

Exemple-2:

El martes último la multinacional de los caramelos Sugus presentó en la Dirección General de Trabajo un expediente de rescisión de contrato para sus 172 trabajadores.

Per definir el moment temporal de la presentació de l'expedient utilitzarem el verb (*presentó*) y l'expressió temporal (*el martes último*).

Pel temps verbal sabem que

$$\begin{aligned} S \text{ Rel } E &> e_i \\ S \text{ Rel } R &> e_i \\ E \text{ Rel } R &= d \text{ c a} \end{aligned}$$

Per la notícia sabem que S és el 9 d'octubre de 1988 i diumenge.

L'expressió temporal és de la classe QUAN-0 i dels comentaris anteriors es dedueix que defineix R.

Per la Base de Coneixement sabem que E ha de ser un punt (moment de la presentació).

De la combinació de totes aquestes dades resulta:

$$\begin{aligned} S \text{ Rel } E &> \\ S \text{ Rel } R &> \\ E \text{ Rel } R &= (\text{única possibilitat entre punts}) \end{aligned}$$

Per càlcul podem obtenir que R és el 4 d'octubre i el temps verbal ens força a que E també sigui el 4 d'octubre de 1988. En aquest cas ha quedat totalment definit allò que ens interessava.

Exemple-3:

El grupo asegurador de capital suizo Winterthur tuvo unos ingresos por recaudación de primas de 8.077 millones en el primer semestre del año.

L'expressió temporal és de la classe QUAN-EN-INTERVAL i acompanya a un verb de resum-gestió. Aquesta situació és de càlcul més senzill doncs considerem directament que ens han definit E (el període de la recaptació).

El sistema ha de calcular l'interval i per tant esbrinar de quin any es tracta. En aquest cas agafa l'any en curs que obté de la data de la notícia (el 30 de setembre de 1986).

El sistema guardarà un resum econòmic de tipus ingressos associat al primer semestre de 1986.

Exemple-4:

Las ventas de la primera compañía papelerá española, Sarrió, alcanzaron en 1986 los 33.046 millones de pesetas, lo que representa un incremento del 6,6% sobre el año anterior.

Per calcular el moment temporal associat a les vendes per valor de 33.046 milions de pessetes, el procés a seguir és paral·lel al descrit per l'exemple 3.

El càlcul del moment associat al període amb el qual es compara s'ha de fer amb precaució. La data de la notícia és el 3 d'abril de 1987. L'expressió *año anterior* no es pot calcular per referència a S doncs obtindriem 1986 i no pas 1985 com informa la notícia. Aquest és un exemple de canvi del referent temporal, que habitualment és la data de la notícia. El sistema ha de recuperar el moment temporal de les vendes i sobre ell aplicar l'expressió *año anterior*. Com a moment temporal comparat es guardarà l'any 1985.

Podem trobar situacions similars on l'interval no agafi tot un any, per exemple es tracta del primer semestre. En aquest cas, calcular el període amb el qual es compara vol dir situar-se en un altre any, però també en el primer semestre.

4.2.5 Conclusions

Al llarg de l'apartat 4.2 hem desenvolupat la problemàtica de l'expressió del temps en el llenguatge. El centre d'atenció han sigut els temps verbals i les expressions temporals.

En quan al temps verbals remarquem que l'assignació de les estructures E-R-S a cadascun d'ells s'ha fet en funció del castellà en general i no particularitzant sobre el domini. El que sí s'ha fet és simplificar-ne alguna si d'aquesta manera s'afavoria l'eficiència del sistema. S'han implementat els temps verbals més freqüents en el domini, però els mecanismes de representació i d'interpretació estan oberts a qualsevol nova incorporació.

En quan a les expressions temporals és evident que hem fet una tipificació d'aquelles que apareixen en el domini, però hem tractat de distingir entre les que tenen

un caràcter general, vàlides per a qualsevol domini, les que són en certa mida generals dins del domini i les més particulars. Al mateix temps, un canvi de domini comportarà una nova tipologia; però, fent l'estudi en la línia del presentat aquí, els mecanismes d'interpretació i de combinació seran totalment aplicables.

5. Components estructurals de COTEM

El sistema té dues components estructurals: la Base de Coneixement i el Sistema Temporal. Un conjunt de procediments, alguns interns i altres externs, serveixen per crear, modificar o eliminar elements dins de cada component. Es pot passar de la Base de Coneixement al Sistema Temporal i a l'inrevés, no es tracta de components aïllades.

El coneixement del domini (empresa, compra, ampliació de capital, tipus de notícies, tipus d'elements temporals...), el coneixement de caràcter general (persona, adreça, caràcter, enter, conjunt...) i el coneixement lingüístic necessari per tractar el domini (*frames* verbals, temps verbals, quans...) estan representats a la Base de Coneixement.

El Sistema Temporal constitueix la representació de la informació temporal. Consta d'una xarxa on els nodes són instàncies de les unitats temporals bàsiques i un calendari on cada any es representa mitjançant un arbre on els nodes són els elements de calendari habituals (dia, quinzena, mes, trimestre, semestre). També inclou tot un conjunt de pseudo-elements de calendari que serveixen de patró per crear els elements de calendari que interessin i només aquells.

Tots els elements d'aquestes dues components s'han implementat com esquemes [Castell-88a] de Knowledge Craft. Aquesta eina ens permet definir jerarquies de diversos tipus (generalització, instanciació...) entre els objectes, associant els mecanismes d'inferència estructural i d'herència de propietats que faciliten i controlen tant la incorporació com la deducció d'informació. També podem establir relacions entre els objectes, tant relacions predefinides (IS-A, PROTOTYPE-OF...) com relacions pròpies de l'usuari (ES-UN, INSTANCIA-DE...). En aquest segon cas podem definir el comportament de les relacions indicant criteris de transitivitat, d'herència, restriccions d'aplicabilitat...

Els objectes són de dos tipus: classes i instàncies. Les classes són descripcions genèriques tant de conceptes primitius (enter, caràcter, string-de-caràcter, element-temporal {punt, interval, cadena}, conjunt, seqüència, seqüència-temporal...) com de conceptes propis de l'aplicació (empresa, compra, persona, inversió...). Les instàncies són realitzacions concretes de les classes, és a dir, les dades de l'aplicació.

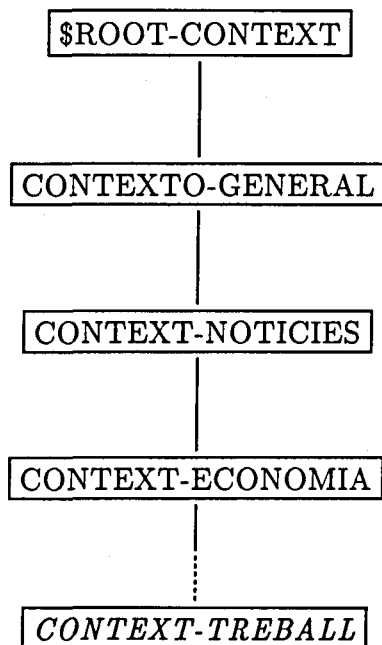
El coneixement sobre un objecte pot expressar-se mitjançant una part declarativa i un conjunt de procediments associats. La part declarativa consisteix en (1) informació sobre la situació de l'objecte dins l'estructura jeràrquica {Superclasses, Subclasses} i (2) un conjunt de parelles descriptor/valor que defineixen l'estructura particular de l'objecte. El valor d'un descriptor és el nom d'una classe (un tipus de valor) quan l'objecte que s'està definint és una descripció genèrica mentre que, si s'està definint una instància, el valor del descriptor és la realització d'aquesta classe (una dada).

Els objectes poden tenir procediments associats a nivell de descriptor i de classe. Els procediments associats a nivell de descriptor serveixen, per exemple, per donar un valor per defecte al descriptor en cas de que no s'hagi fet cap assignació prèvia o per controlar que el valor assignat compleixi unes determinades condicions. Per associar procediments a nivell de classe s'ha utilitzat el concepte de mètode de Knowledge Craft. Per exemple la classe que defineix les seqüències té associats diferents mètodes per realitzar les operacions típiques de tractament de seqüències com són: obtenir el primer element, obtenir el següent element, buidar la seqüència, consultar si estem a principi o final de seqüència, etc.

També les relacions poden tenir procediments associats: els *demons*. Amb aquest procediments podem comprovar condicions complexes abans de permetre la incorporació d'una nova relació entre dos esquemes, podem provocar la creació d'altres relacions com efecte colateral, etc.

El Knowledge Craft permet agrupar els esquemes per contextos y establir una jerarquia de contextos. Hem utilitzat aquesta possibilitat d'organització de cara a definir una representació modular que pugui ser aprofitada per altres aplicacions.

Així doncs, el sistema treballa amb la següent jerarquia de contextos.



En el primer nivell, \$ROOT-CONTEXT, estan definits tots els esquemes bàsics de Knowledge Craft. És el context predefinit. Nosaltres no hi hem afegit res.

En el segon nivell, CONTEXTTO-GENERAL, estan definits els conceptes de

coneixement general. Per exemple: “clase”, “instancia”, “descriptor”, “secuencia”, “grafo”... i també relacions com “es-un”, “instancia-de”, “miembro-de”...

En el tercer nivell, CONTEXT-NOTICIES, estan definits els conceptes considerats com a generals per a qualsevol tipus de notícia. La jerarquia d'elements temporals i la dels elements de calendari està definida en aquest nivell, així com les relacions temporals. També s'inclou en aquest context una jerarquia d'esquemes auxiliars necessaris pel processament de les notícies i que seran especialitzats a cada subcontext en funció del domini de les notícies. Per exemple conté la descripció general dels *frames* verbals, dels temps verbals, de les expressions temporals...

En el quart nivell, CONTEXT-ECONOMIA, estan definits tots els conceptes propis de les notícies d'economia. Des de les especialitzacions dels conceptes del context superior (els *frames* verbals amb el seu significat en aquest context, els tipus de notícies, les components d'una notícia) fins els conceptes econòmics pròpiament dits (empresa, ampliació-de-capital, beneficis, compra, tancament...). També en aquest nivell tindrem totes les instàncies: les dades del sistema.

El cinquè nivell, CONTEXT-TREBALL, és un context auxiliar que s'utilitza durant el procés de tractament de la notícia i que és eliminat un cop incorporada la informació a la Base de Coneixement i al Sistema Temporal. Aquí tindrem instàncies dels *frames* verbals, dels *quans*, de les components estructurals de les notícies i, també, de conceptes propis del domini econòmic i d'elements del Sistema Temporal.

A títol il·lustratiu presentem alguns dels esquemes utilitzats per COTEM. Uns esquemes contenen la relació IS-A mentre la majoria contenen la relació ES-UN. La relació IS-A és la relació predefinida a Knowledge Craft per tal d'establir jerarquies de conceptes. La relació ES-UN és similar a IS-A amb la diferència de que l'hem definida nosaltres i per tant podem modificar-ne el funcionament canviant criteris de transitivitat, d'herència... Sempre que es pot s'utilitza la relació ES-UN, però quan s'han de fer servir facilitats de l'eina hem d'explicitar la tipologia de l'esquema amb la relació IS-A.

El primer esquema es tracta d'una classe que descriu una estructura de dades: la seqüència. És un concepte general i com a tal està situat en el CONTEXTO-GENERAL. Té definits descriptors que són mètodes i que justament implementen les operacions sobre seqüències: obtenir el primer element de la seqüència (“primero”), passar al següent (“siguiente”), afegir un element (“anadir-elem”)... Els descriptors “recorrida” i “en-curso” contenen en tot moment els elements de la seqüència repartits entre els ja consultats i els pendants.

Un parell de descripcions de mètodes apareixen a continuació d'aquest esquema.


```

{{ SECUENCIA
  ES-UN: ESTRUCTURA-BASICA
  TIENE-POR-SUBCLASE-A: RESULTATS-ANALITZADOR
  RECORRIDA:
  EN-CURSO:
  VACIAR: VACIAR
  PRIMERO: PRIMERO
  SIGUIENTE: SIGUIENTE
  ANTERIOR: ANTERIOR
  ES-FINAL: ES-FINAL
  PRINCIPIO: PRINCIPIO
  BORRAR-ELEM: BORRAR-ELEM
  ANADIR-ELEM: ANADIR-ELEM
  CUANTOS-ELEM: CUANTOS-ELEM}}

```

```

{{ VACIAR
  IS-A: METHOD}}

```

```

{{ SIGUIENTE
  IS-A: METHOD}}

```

La descripció dels *quans* es situa a nivell de CONTEXT-NOTICIES. Presentem l'esquema més general "quan" on consten totes les subclasses definides. Posseeix dos descriptors que cada subclasse pot heretar si l'interessa. En particular, el descriptor "convertir-quan" és heretat per totes les subclasses doncs és el mètode que cal activar per convertir en element temporal una expressió temporal. De les seves subclasses es crearan instàncies durant el procés d'interpretació de les notícies. Aquestes instàncies només existiran en el CONTEXT-TREBALL i desapareixeran un cop incorporada la notícia al sistema.

```

{{ QUAN
  ES-UN: ESQ-AUX-TEMPORAL
  TIENE-POR-SUBCLASE-A:
    QUAN-VA-DE-UNITAT QUAN-SUBINTERVAL QUAN-PREVISIO
    QUAN-SITUADOR-INDEFINIT QUAN-GERMA QUAN-FINS QUAN-FINAL
    QUAN-FA-CONCRET QUAN-EN-LAPSE-CONCRET QUAN-EN-INTERVAL
    QUAN-DES-DE-FINS QUAN-DES-DE QUAN-DE QUAN-A-DIRIGIT QUAN-O
  INTRODUTOR:
  CONVERTIR-QUAN:}}

```

El “quan-especific” és un esquema auxiliar que, tal com s’explica en el descriptor “descripcion”, representa les expressions temporals més específiques. Moltes subclasses de “quan” contenen un o més descriptors tals que el seu valor ha de ser una instància d’aquest esquema.

De fet, aquest esquema representa les expressions recollides per l’arbre d’anàlisi QUAN-0 (veure capítol 4). El descriptor “punt-inter” serveix per indicar si l’expressió es converteix en un punt o en un interval. Els quatre següents descriptors recullen el que s’anomena unitat temporal explícita i els tres últims representen el marc temporal. Els descriptors que contenen la paraula “tipus” tenen com a valor un nom d’element de calendari (dia, mes, any...). Els que contenen la paraula “valor” recullen números de dies o anys, noms de mesos... El descriptor “posicio” és un indicador de present, passat o futur. El descriptor “posic-ref” afegeix a aquestes tres la possibilitat d’indicar un any. El descriptor “ordre” conté valors numèrics en el cas d’expressions que tracten de trimestres (1, 2, 3, 4), semestres (1, 2) o quinzenes (1, 2).

```
{ { QUAN-ESPECIFIC
  ES-UN: ESQ-AUX-TEMPORAL
  PUNT-INTER:
  TIPUS-EXPL:
  VALOR-EXPL:
  POSICIO:
  ORDRE:
  TIPUS-REF:
  VALOR-REF:
  POSIC-REF:
  DESCRIPCION: "Amb aquest esquema representem les expressions
                temporals mes especificques, les que ens marquen un
                element de calendari. Es un esquema nucli per a
                diferents quans" }
```

El “quan-de” és una subclasse de “quan” que representa les expressions del mateix nom ja descrites en el capítol anterior. Remarquem que el descriptor “convertir-quan” té valor: “met-quan-de”, el mètode per convertir a elements temporals les expressions d’aquest tipus. Posseeix un altre descriptor que és un mètode, “calcul-expressio”, que està associat al descriptor “expressio” i que serveix per obtenir el nom de l’element de calendari que correspon a l’expressió temporal continguda a “expressio”.

```

{{ QUAN-DE
  ES-UN: QUAN
  EXPRESSIO:
  DESCRIPCION: "Aquest quan es identic al QUAN-0 pel que fa
                a l'expressio temporal. Ara be, te la particularitat
                de que esta perfectament definit a qui cal associar
                aquesta informacio temporal: al grup nominal anterior
                a la preposicio DE (reunion, beneficios...)"
  INTRODUTOR: "DE"
  CALCUL-EXPRESSIO: MET-ELEM-CONCRET
  CONVERTIR-QUAN: MET-QUAN-DE}}

```

L'esquema següent descriu un descriptor. El Knowledge Craft permet restringir els esquemes en els quals pot aparèixer un descriptor ("domain") i també quins són els seus valors possibles ("range"). Per tant "expressio" és un descriptor vàlid només per algunes subclasses de "quan" i ha de tenir com a valor una instància de "quan-especific".

```

{{ EXPRESSIO
  ES-UN: DESCRIPTOR
  DOMAIN:
    (OR (TYPE ES-UN QUAN-0) (TYPE ES-UN QUAN-DE)
        (TYPE ES-UN QUAN-PREVISIO) (TYPE ES-UN QUAN-A-DIRIGIT))
  RANGE: (SCHEMA (TYPE INSTANCIA-DE QUAN-ESPECIFIC))}}

```

Els tres esquemes següents mostren com es representen els patrons dels temps verbals. L'esquema genèric és "temps-verbal" que té una subclasse per cadascun dels temps verbals més freqüents del domini. El descriptor "adaptar" és un mètode que permet especialitzar les relacions genèriques (descriptors "relacions-X-Y") en funció del tipus d'elements temporals a relacionar.

```

{{ TEMPS-VERBAL
  ES-UN: ESQ-AUX-TEMPORAL
  TIENE-POR-SUBCLASE-A:
    T-FUTUR-PERFET T-FUTUR-IMPERFET T-INDEFINIT
    T-IMPERFET T-PRET-PERFET T-PRESENT
  ADAPTAR: ADAPTAR
  RELACIONS-E-R:
  RELACIONS-E-S:
  RELACIONS-S-R:}}

```

```

{{ T-PRESENT
  ES-UN: TEMPS-VERBAL
  RELACIONS-E-S:
    IGUAL DESPRES ENLLACAT-PER CONTE
    COMENCAT-PER ACABAT-PER
  RELACIONS-E-R:
    IGUAL DURANT COMENCA ACABA
  RELACIONS-S-R:
    IGUAL ABANS ENLLACA DURANT
    COMENCA ACABA}}

```

```

{{ ADAPTAR
  IS-A: METHOD}}

```

L'esquema "frame-verbal" és la descripció genèrica dels verbs amb tots els casos comuns. Tots els descriptors que comencen amb la paraula "quan" han de contenir instàncies d'alguna subclasse de "quan".

```

{{ FRAME-VERBAL
  ES-UN: ESQUEMA-AUXILIAR
  TIENE-POR-SUBCLASE-A:
    FV-OPERADOR FV-DESCRIPCIO
    FV-ACCIO ELEMENT-ESTRUC-BASIC
  AGENT:
  TEMA:
  EXPERIMENTADOR:
  BENEFICIARI:
  CAS-QUAN:
  QUAN-AGENT:
  QUAN-TEMA:
  QUAN-EXPERIMENTADOR:
  QUAN-BENEFICIARI:
  INSTRUMENT:
  TRIPLETA-VERBAL:
  PREVISIO: NIL
  EN-LLOC:
  DES-DE-LLOC:
  FINS-LLOC:

```

EN-VALOR:
DES-DE-VALOR:
FINS-VALOR:}}

Tres esquemes més mostren part de la representació de la tipologia verbal que s'ha fet pel domini, els mètodes associats als *frames* verbals i un verb concret amb els seus casos particulars i els seus mètodes. Tots aquests esquemes ja pertanyen a CONTEXT-ECONOMIA.

```
{{ FV-GESTIO
  ES-UN: FV-ACCIO
  TIENE-POR-SUBCLASE-A:
    FV-TANCAMENT FV-RESUM-GESTIO FV-INTEGRACIO
    FV-FUSIO FV-CREACIO FV-COMPRA-VENTA
    FV-COMANDA FV-AMPLIACIO-DE-CAPITAL
  ACTIVA-SUCCESES:
  REGLES-CREACIO:
  REGLES-COMPLETESA:}}
```

```
{{ REGLES-CREACIO
  IS-A: METHOD}}
```

```
{{ REGLES-COMPLETESA
  IS-A: METHOD}}
```

```
{{ AMPLIAR-1
  ES-UN: FV-AMPLIACIO-DE-CAPITAL
  TEMA: CAPITAL-SOCIAL
  CARREGADA-A:
  PROPORCIO:
  COM:
  MOTIU:
  REGLES-CREACIO: CREAR-SUCCESES-AMPLIAR-1
  REGLES-COMPLETESA: COMPLETAR-SUCCESES-AMPLIAR-1}}
```

Per últim presentem alguns esquemes del CONTEXT-ECONOMIA i que corresponen a conceptes propis del domini. Hi ha un concepte general, "succes", que té dues subclasses: "estat" i "fet". Les subclasses que ens interessin ara de cadascun d'aquests conceptes són "resum-economic" i "fet-economic". Presentem la descripció i un exemple del primer, seguits d'un exemple del segon. Tot els successos poden ser marcats (descriptor "real-previsio") com reals o com previstos segons es dedueixi de les notícies.

Aquests esquemes són els que lliguen la Base de Coneixement amb el Sistema Temporal doncs tenen descriptors que el seu valor ha de ser una instància d'element temporal. És el cas dels descriptors "moment-decisio", "moment-proposta", "moment-ratificacio" i "moment-aplicacio" de l'esquema "ampliacio-capital". I també "moment-anterior" de "resum-economic" i "moment-pagament-dividends" i "moment-proposta" de "pagament-dividends".

{{ RESUM-ECONOMIC

TIENE-POR-SUBCLASE-A:

VENDES TOTAL-CAPITAL RESERVES PERDUES

PAGAMENT-DIVIDENDS INVERSIONS INGRESOS

FACTURACIONES CREACIONES DESPESES COMPRES

CASH-FLOW BENEFICIS BALANC AMORTITZACIONES

ES-UN: ESTAT

SEGUENT-RESUM: 1

QUANTITAT-RESUM:

REAL-PREVISIO:

TPC-RESPECTE-ANTERIOR:

MOMENT-ANTERIOR:}}

{{ PAGAMENT-DIVIDENDS

ES-UN: RESUM-ECONOMIC

MOMENT-PAGAMENT-DIVIDENDS:

PER-TIPUS-ACCIO:

PROPOSADA-A:

MOMENT-PROPOSTA:}}

{{ AMPLIACIO-CAPITAL

ES-UN: FET-ECONOMIC

INTERESSAT:

DECIDIDA-PER:

MOMENT-DECISIO:

LLOC-DECISIO:

PROPOSADA-A:

MOMENT-PROPOSTA:

REBUTJADA-PER:

MOMENT-RATIFICACIO:

LLOC-RATIFICACIO:

IMPORT-AMPLIACIO:

PERCENTATGE:
NOMBRE-ACCIONS:
VALOR-ACCIO:
PAS-ACCIO:
CARREGADA-A:
COST-ACCIONISTA:
PARTICIPACIO-RESULTATS:
NOU-TOTAL-CAPITAL:
MOMENT-APLICACIO:}}

6. El Sistema Temporal

Ja hem exposat el tipus d'informació temporal que apareix a les notícies i com s'expressa la noció del temps en el llenguatge, presentem ara el Sistema Temporal utilitzat per representar aquesta informació [Castell-87]. Cal tenir en compte que la Base de Coneixement i el Sistema Temporal estan totalment relacionats, qualsevol informació relativa al temps associada al coneixement es representa en el sistema temporal mentre que a la Base de Coneixement tant sols hi ha una referència a la representació corresponent. D'altra banda, si es demana quins fets s'han produït en un determinat interval de temps, a partir del Sistema Temporal accedirem a la Base de Coneixement per obtenir-ne la resposta.

Al sistema se li fan dos requeriments: poder expressiu, ha de permetre la representació de les manifestacions temporals del llenguatge, i eficiència tant en els mecanismes de raonament com de propagació de relacions.

6.1 Components del Sistema Temporal

El Sistema Temporal consta, ja ho hem mencionat, d'una xarxa on els nodes són instàncies de les unitats temporals bàsiques i un calendari on cada any es representa mitjançant un arbre. La informació temporal s'associa als objectes de la Base de Coneixement creant lligams entre ells i els nodes de la xarxa temporal (figura 6.1). Si les instàncies de la xarxa fan referència a informació temporal específica (un dia, un mes, el primer trimestre d'un any...), llavors tenen associats els elements corresponents en el calendari i les denominem instàncies tancades. Els nodes de la xarxa que no tenen informació temporal específica s'anomenen instàncies obertes.

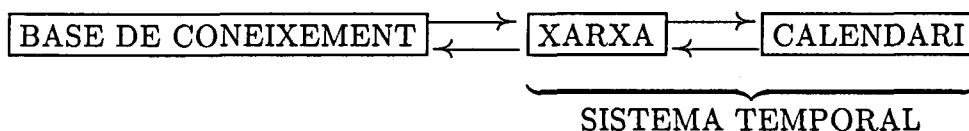


Figura 6.1

El fet de que el sistema temporal consti de dos elements (xarxa i calendari) distingeix aquest model de representació dels treballs publicats fins ara. Així doncs, quan dues instàncies temporals tenen totes dues elements del calendari associats, no es manté cap relació entre elles explícitament a la xarxa. Hom considera que el càlcul de relacions entre elements del calendari és fàcil i poc costós, per tant només es mantenen explícites les relacions que no es poden obtenir a través del calendari i que són presents a la notícia tractada.

6.1.1 La Xarxa

La jerarquia d'unitats temporals que COTEM té definida actualment queda reflectida a la figura 6.2:

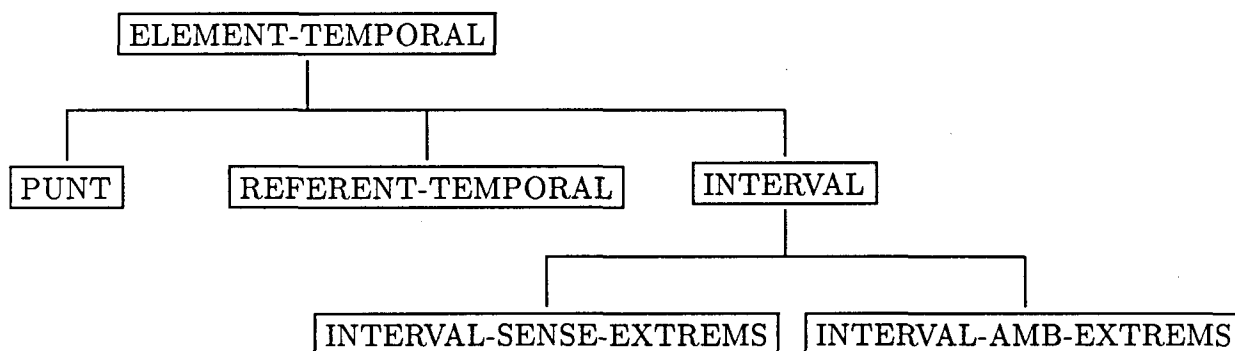


Figura 6.2

Essent la descripció d'aquests esquemes la següent:

```
{{ ELEMENT-TEMPORAL
TIENE-POR-SUBCLASE-A: REFERENT-TEMPORAL PUNT INTERVAL
ES-UN: CLASE
TIPUS:
RELACIONAR: RELACIONAR
RELACIONAT-AMB:
RELACIONS-MANTINGUDES:}}
```

```
{{ PUNT
ES-UN: ELEMENT-TEMPORAL
VALOR:}}
```

```
{{ INTERVAL
ES-UN: ELEMENT-TEMPORAL
TIENE-POR-SUBCLASE-A:
    INTERVAL-SENSE-EXTREMS INTERVAL-AMB-EXTREMS
LONGITUD:}}
```

```
{{ INTERVAL-SENSE-EXTREMS
ES-UN: INTERVAL
VALOR:}}
```

```

{{ INTERVAL-AMB-EXTREMS
  ES-UN: INTERVAL
  INICI:
  FINAL:}}

{{ REFERENT-TEMPORAL
  ES-UN: ELEMENT-TEMPORAL
  TIPUS: OBERT}}

{{ TIPUS
  ES-UN: DESCRIPTOR
  RANGE: (OR OBERT TANCAT)
  DOMAIN: (TYPE ES-UN ELEMENT-TEMPORAL)
  DESCRIPCION:
    "Descriptor OBLIGATORI per tot element-temporal i que
    serveix per indicar si te associat [tancat] o no [obert]
    un element del calendari"}}

{{ VALOR
  ES-UN: DESCRIPTOR
  DOMAIN: (OR (TYPE ES-UN INTERVAL-SENSE-EXTREMS) (TYPE ES-UN PUNT))
  RANGE: (SCHEMA (TYPE INSTANCIA-DE ELEMENT-CALENDARI))}}

{{ INICI
  ES-UN: DESCRIPTOR
  DOMAIN: (TYPE ES-UN INTERVAL-AMB-EXTREMS)
  RANGE: (SCHEMA (TYPE INSTANCIA-DE ELEMENT-CALENDARI))}}

{{ FINAL
  ES-UN: DESCRIPTOR
  DOMAIN: (TYPE ES-UN INTERVAL-AMB-EXTREMS)
  RANGE: (SCHEMA (TYPE INSTANCIA-DE ELEMENT-CALENDARI))}}

```

Les instàncies d'aquests conceptes bàsics constitueixen la xarxa temporal que representa la informació temporal associada al coneixement econòmic en cada moment. Els elements de la Base de Coneixement tindran aquestes instàncies com a valors dels seus descriptors temporals. Per la seva banda, les instàncies indiquen de quins descriptors són valor.

Com es pot veure a la presentació anterior, tot element temporal té els descriptors "tipus", "relacionat-amb", "relacions-mantingudes" i "relacionar". Els punts tenen un descriptor addicional, "valor", que només podem omplir amb la referència a un

element de calendari. Els intervals tenen el descriptor addicional "longitud" per indicar el nombre de dies de que consten.

S'observa que hi ha dos tipus d'intervals: intervals amb extrems i intervals sense extrems. Aquests últims són intervals que corresponen directament a períodes de calendari com any, trimestre, mes... i per tant tenen un descriptor addicional, també anomenat "valor", que cal omplir amb la referència d'un element de calendari. Els intervals amb extrems són aquells que es defineixen utilitzant dues referències al calendari, no cal que els extrems siguin dies. Aquest segon tipus d'interval té dos descriptors addicionals, "inici" i "final", que s'omplen amb referències a elements de calendari de qualsevol tipus (sempre que siguin períodes).

El "tipus" és un descriptor obligatori i té dos valors possibles: "obert" o "tancat". Els elements temporals que fan referència a elements de calendari tindran aquest descriptor amb valor "tancat" mentre que els elements temporals no definits tindran el valor "obert". Per tant els intervals sense extrems i els punts tancats han de tenir ple el descriptor "valor" mentre que els intervals amb extrems, tancats, han de tenir plens els descriptors "inici" i "final".

El "referent-temporal" és un tipus d'element temporal que s'utilitza quan es desconeix si la informació temporal que representa és punt o interval. És evident que ha de tenir obligatòriament el descriptor "tipus" amb valor "obert". Com el sistema actual manté relacions especialitzades en funció dels tipus d'elements relacionats, les instàncies de "referent-temporal" no intervenen en el procés de raonament del Sistema Temporal. Val a dir, que desconèixer el tipus d'element temporal és una situació poc freqüent.

Entre els elements de la xarxa es mantenen les relacions temporals definides per Allen [Allen-83], però especialitzades segons els tipus d'elements que connecten. Les relacions que manté el sistema són:

Codi numèric	Codi simbòlic	Nom complet	Codi de la relació inversa
0	<	abans	1
1	>	després	0
2	D	durant	3
3	Di	conté	2
4	S	solapa	5
5	Si	solapat per	4
6	E	enllaça	7
7	Ei	enllaçat per	6
8	C	comença	9
9	Ci	començat per	8
10	A	acaba	11
11	Ai	acabat per	10
12	=	igual	12
13	PAI	punt abans interval	21
14	PDI	punt després interval	20
15	PEI	punt enllaça interval	23
16	PEPI	punt enllaçat per interval	22
17	PCI	punt comença interval	24
18	PFI	punt finalitza interval	25
19	PCPI	punt durant interval	26
20	IAP	interval abans punt	14
21	IDP	interval després punt	13
22	IEP	interval enllaça punt	16
23	IEPP	interval enllaçat per punt	15
24	ICPP	interval començat per punt	17
25	IFP	interval finalitzat per punt	18
26	ICP	interval conté punt	19
27	PAP	punt abans punt	28
28	PDP	punt després punt	27
29	PIP	punt igual punt	29
30	PEP	punt enllaça punt	31
31	PEPP	punt enllaçat per punt	30

Les tretze primeres corresponen a les relacions entre intervals definides per Allen.

Pot ser que ens manqui informació per delimitar exactament la relació que hi ha entre dos nodes, llavors mantenim una relació múltiple i confiem en dades posteriors que permetin concretar-la més. Atesa aquesta multiplicitat, no és possible utilitzar el concepte de relació que té predefinit el Knowledge Craft, doncs les relacions múltiples es mantenen de forma independent. Aixó fa molt difícil el procés de propagació de relacions que és un dels trets bàsics del Sistema Temporal.

Per tant les relacions temporals són gestionades per COTEM. Cada element temporal té un descriptor, "relacionat-amb", on consta la llista de tots els elements temporals amb els quals està relacionat de manera explícita. En el descriptor "relacions-mantingudes" hi ha una altra llista relacionada biunívocament amb l'anterior. Cada element d'aquesta segona subllista és una parella de valors: un primer valor, la llista de relacions temporals mantingudes entre el propietari del descriptor i l'element temporal corresponent de la llista associada i, un segon valor, un indicador de si aquesta informació s'ha tret directament del text o per propagació. El descriptor "relacionar" conté el mètode que provoca la propagació de relacions quan volem afegir una nova relació temporal.

6.1.2 El Calendari

La jerarquia d'elements de calendari que COTEM té definida actualment queda reflectida a la figura 6.3:

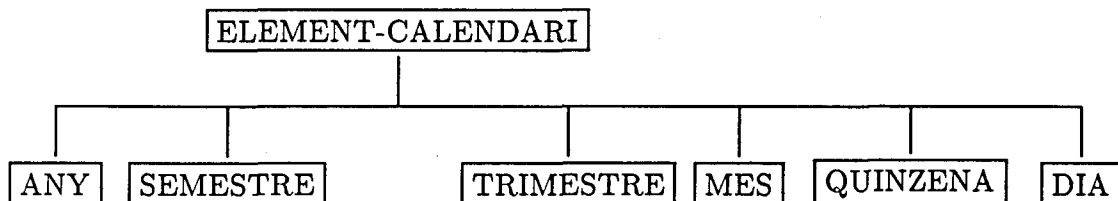


Figura 6.3

Essent la descripció d'aquests esquemes la següent:

```

{{ ELEMENT-CALENDARI
  TIENE-POR-SUBCLASE-A: TRIMESTRE SEMESTRE QUINZENA MES DIA ANY
  ES-UN: CLASE
  PARE-DE:
  GERMA-DRET-VIA:
  GERMA-ESQUERRE-VIA:
  FILL-VIA:}}
  
```

```

{{ ANY
  ES-UN: ELEMENT-CALENDARI
  TIENE-POR-INSTANCIA-A:}}

{{ SEMESTRE
  ES-UN: ELEMENT-CALENDARI
  TIENE-POR-INSTANCIA-A:}}

{{ TRIMESTRE
  ES-UN: ELEMENT-CALENDARI
  TIENE-POR-INSTANCIA-A:}}

{{ MES
  ES-UN: ELEMENT-CALENDARI
  TIENE-POR-INSTANCIA-A:}}

{{ QUINZENA
  ES-UN: ELEMENT-CALENDARI
  TIENE-POR-INSTANCIA-A:}}

{{ DIA
  ES-UN: ELEMENT-CALENDARI
  TIENE-POR-INSTANCIA-A:}}

{{ PARE-DE
  ES-UN: DESCRIPTOR
  DOMAIN: (AND (TYPE ES-UN ELEMENT-CALENDARI)
              (NOT (TYPE ES-UN DIA)))
  CARDINALITY: (0 15)}}

{{ FILL-VIA
  ES-UN: DESCRIPTOR
  DOMAIN: (AND (TYPE ES-UN ELEMENT-CALENDARI)
              (NOT (TYPE ES-UN ANY)))
  RANGE: (OR COMENCA ACABA DURANT)
  CARDINALITY: (0 1)}}

{{ GERMA-DRET-VIA
  ES-UN: DESCRIPTOR
  DOMAIN: (TYPE ES-UN ELEMENT-CALENDARI)
  RANGE: (OR ABANS ENLLACA)
  CARDINALITY: (0 1)}}

```

```

{{ GERMA-ESQUERRE-VIA
  ES-UN: DESCRIPTOR
  CARDINALITY: (0 1)
  DOMAIN: (TYPE ES-UN ELEMENT-CALENDARI)
  RANGE: (OR DESPRES ENLLACAT-PER)}}

```

Entre els elements de calendari també es mantenen relacions temporals, si bé aquestes relacions no es treuen del text sinó que són relacions temporals establertes d'avantmà. De fet, són aquestes relacions les que ens permeten mantenir desconnectats els elements temporals tancats. Donats dos elements temporals tancats, consultarem a quins elements de calendari fan referència, agafarem la relació que hi ha entre ells i deduirem la relació existent entre els elements temporals. Ara bé, entre els elements de calendari les relacions temporals mai són múltiples. Per tant en el calendari sí s'utilitzen relacions definides segons Knowledge Craft. Hem definit les següents relacions:

```

{{ ABANS
  INSTANCE: RELACIO-TEMPORAL-CALENDARI
  INVERSE: DESPRES
  TRANSITIVITY: (OR (REPEAT (STEP ABANS T) 1 INF)
    (LIST (OR (STEP ABANS T) (STEP ENLLACA T)
      (PATH COMENCA) (PATH DURANT))
      (REPEAT (OR (STEP ABANS T) (STEP ENLLACA T)) 1 INF))
    (LIST (PATH ACABA) (PATH ABANS))))}}

```

```

{{ DESPRES
  INSTANCE: RELACIO-TEMPORAL-CALENDARI
  INVERSE: ABANS
  TRANSITIVITY: (OR (REPEAT (STEP DESPRES T) 1 INF)
    (LIST (OR (STEP DESPRES T) (STEP ENLLACAT-PER T)
      (PATH DURANT) (PATH ACABA))
      (REPEAT (OR (STEP DESPRES T) (STEP ENLLACAT-PER T)) 1 INF))
    (LIST (PATH COMENCA) (PATH DESPRES))))}}

```

```

{{ COMENCA
  INSTANCE: RELACIO-TEMPORAL-CALENDARI
  INVERSE: COMENCAT-PER
  TRANSITIVITY: (REPEAT (STEP COMENCA T) 1 INF)}}

```

```

{{ ACABA
  INSTANCE: RELACIO-TEMPORAL-CALENDARI
  INVERSE: ACABAT-PER
  TRANSITIVITY: (REPEAT (STEP ACABA T) 1 INF)}}

{{ DURANT
  INSTANCE: RELACIO-TEMPORAL-CALENDARI
  INVERSE: CONTE
  TRANSITIVITY: (OR (REPEAT (STEP DURANT T) 1 INF)
    (LIST (STEP DURANT T)
      (OR (PATH COMENCA) (PATH DURANT) (PATH ACABA)))
    (LIST (PATH COMENCA) (OR (PATH DURANT) (PATH ACABA)))
    (LIST (PATH ACABA) (OR (PATH DURANT) (PATH COMENCA)))))}}

{{ ENLLACA
  INSTANCE: RELACIO-TEMPORAL-CALENDARI
  INVERSE: ENLLACAT-PER
  TRANSITIVITY:
    (LIST (REPEAT (PATH ACABA) 0 1) (STEP ENLLACA T))}}

{{ ENLLACAT-PER
  INSTANCE: RELACIO-TEMPORAL-CALENDARI
  INVERSE: ENLLACA
  TRANSITIVITY:
    (LIST (REPEAT (PATH COMENCA) 0 1) (STEP ENLLACAT-PER T))}}

```

Cal observar que no s'ha definit la relació solapa ni la seva inversa doncs aquestes relacions no són possibles entre els elements de calendari. També fem notar que les relacions no han sigut especialitzades en funció de si els elements són intervals o punts; d'aquesta manera hom aprofita les regles de transitivitat, que són iguals en tots dos casos, i es fa més fàcil la consulta del calendari. Quan aquestes relacions s'han de transmetre a elements temporals, es fa la corresponent especialització.

Els elements de calendari no estan relacionats tots amb tots. D'entrada només els que pertanyen al mateix any són susceptibles d'estar relacionats, les relacions entre elements de diferent any es trobarà a partir dels pares-any que sí estan relacionats. Dins de l'arbre d'un any, els elements estan relacionats amb el seu pare, el seu germà dret (si existeix) i el seu germà esquerre (si existeix). Per facilitar l'accés des d'un element a l'altre, existeixen els descriptors "germa-dret-via", "germa-esquerre-via" i "fill-via" que contenen el nom de la relació corresponent. El descriptor "pare-de" és informatiu i serveix a l'hora de crear nous fills d'un element de calendari.

Quan es crea un element temporal tancat o quan es vol tancar un element obert, és obligatori que el o els elements de calendari ja existeixin per tal de poder fer l'assignació al descriptor "valor" o als descriptors "inici" i "final". Al sistema no l'interessa mantenir tots els elements de calendari possibles dins del lapse temporal en el qual considerem situada tota la informació que tractarà, a més aquests elements s'han de mantenir connectats per relacions temporals.

Hem optat per la creació dels elements de calendari a mida que es necessiten, mantenint l'estructura mínima de calendari per tal de que el sistema de relacions funcioni. Això representa economia d'espai i també de temps al mantenir tan sols les relacions imprescindibles. Per aquest motiu el Sistema Temporal consta d'un patró (veure figura 6.4 a la pàgina següent) que defineix de manera genèrica la disposició dels elements dins de la jerarquia de l'arbre i les relacions que hi ha entre ells.

Inicialment el sistema només té definits elements de calendari que representen els anys del lapse temporal que tractarà. La situació inicial està expressada gràficament a la figura 6.5.

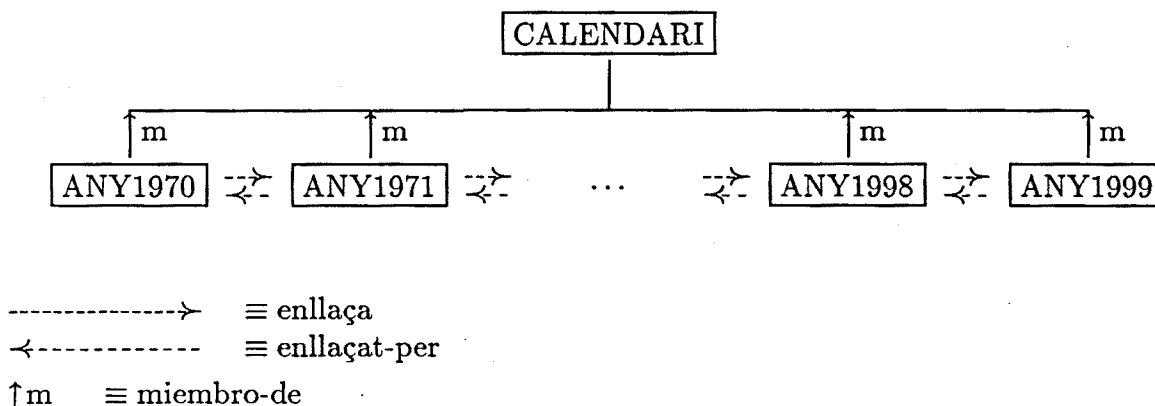


Figura 6.5

Quan interessa un element de calendari, es comprova si existeix o no. En cas de que no existeixi, es crea l'element en qüestió més tots els elements superiors dins de l'arbre i que tampoc existeixin, fins arribar a l'any. Les relacions verticals es treuen del patró directament. Les relacions horitzontals es calculen en el moment de creació, encara que també estan en el patró (no en el nivell dels dies), doncs la relació és diferent segons si el germà és l'immediat o no ho és. D'aquesta manera, en tot moment els elements de calendari creats i pertanyents a un mateix any estan lligats horitzontalment i verticalment. A la figura 6.6 podem veure una possible situació del calendari un cop processades algunes notícies.

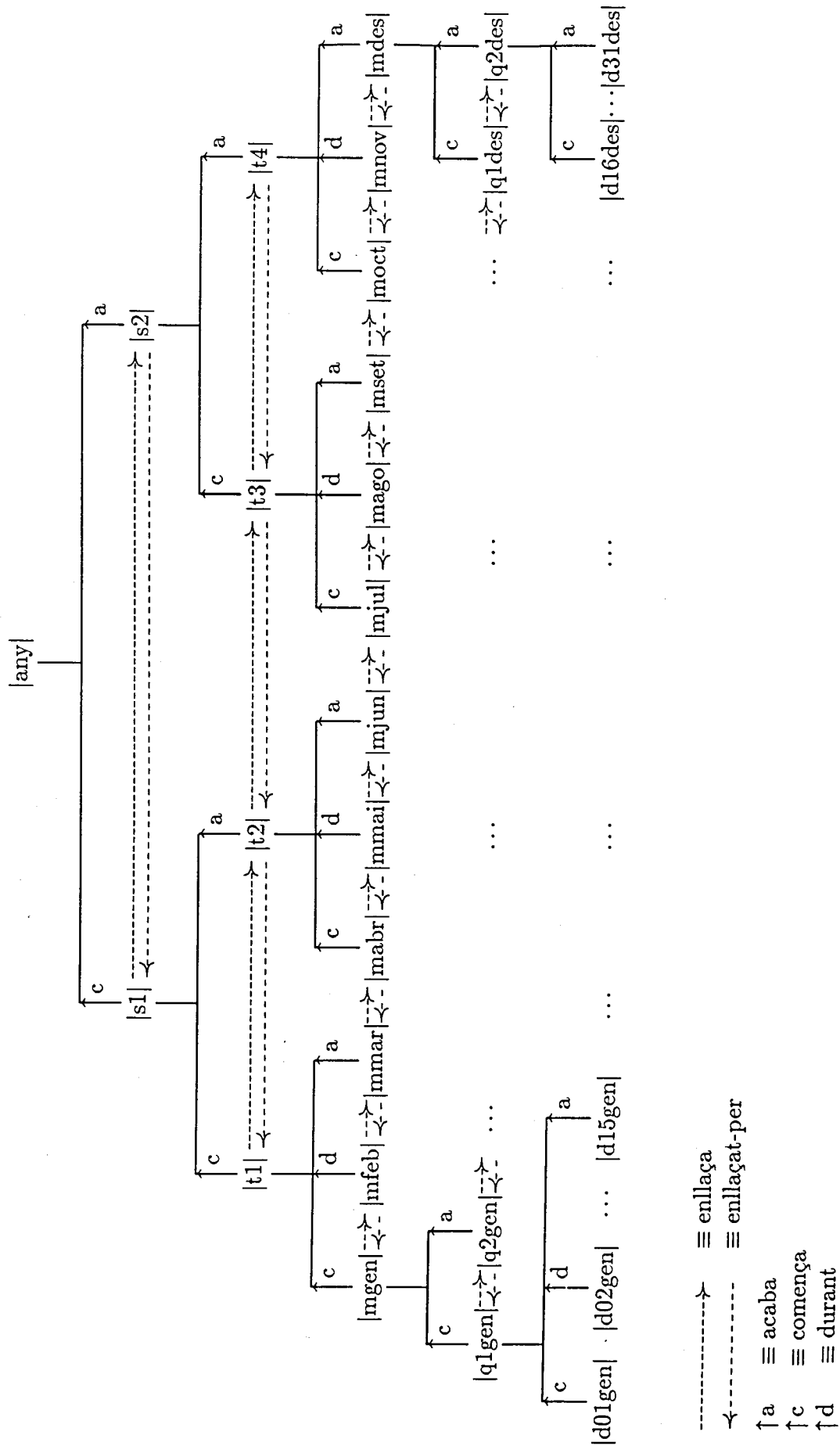
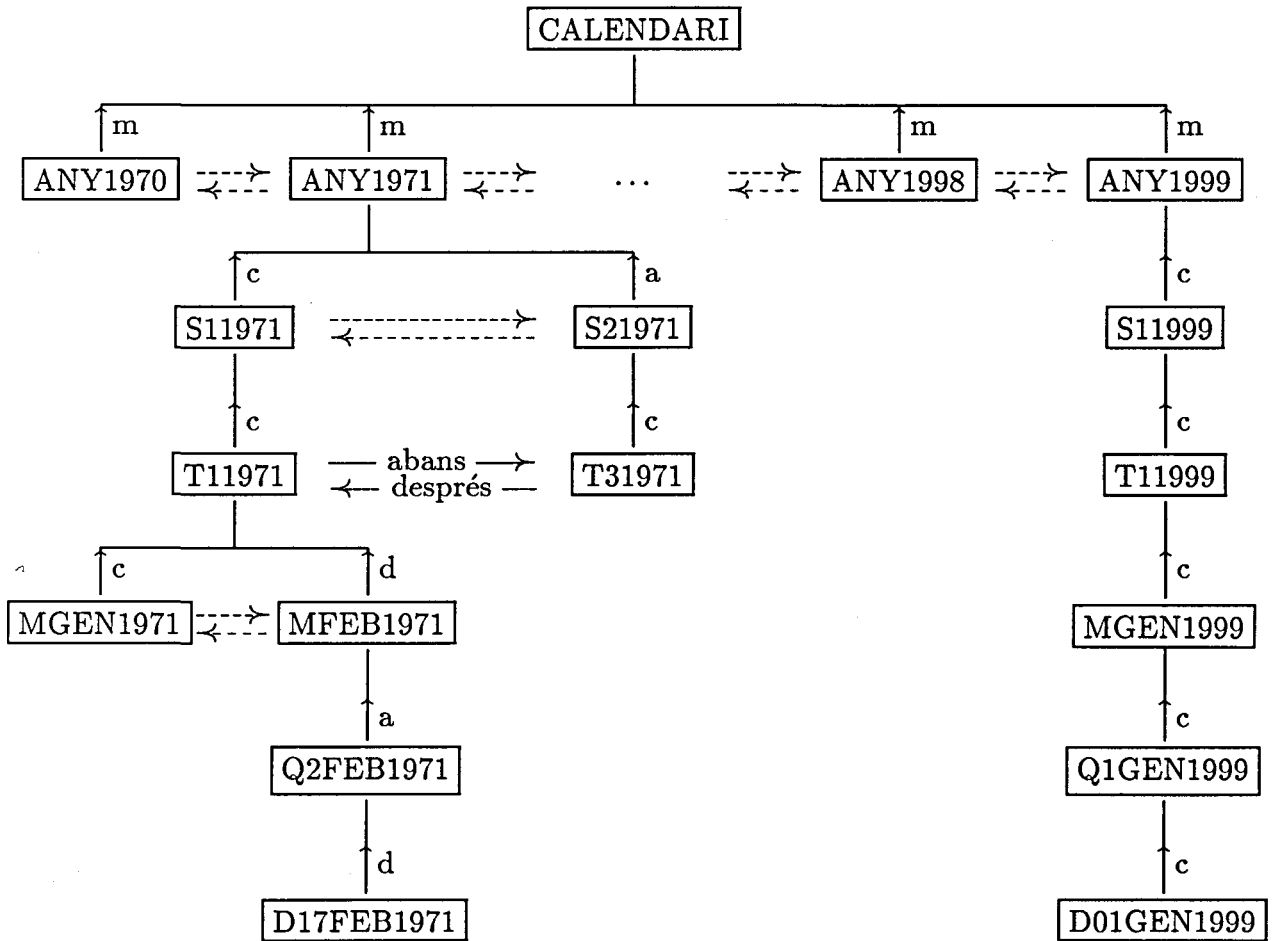


Figura 6.4. Any patró



- > ≡ enllaça
- <----- ≡ enllaçat-per
- ↑a ≡ acaba
- ↑c ≡ comença
- ↑d ≡ durant
- ↑m ≡ miembro-de

Figura 6.6

6.2 Creació, propagació i consulta de relacions temporals

La creació d'una instància d'element temporal activa l'algorisme d'actualització de la xarxa. Si està relacionada amb altres instàncies, caldrà incorporar la relació i veure si això modifica la resta del Sistema Temporal. Si la nova instància és de tipus tancat i hem trobat una relació amb una altra instància tancada, no afegirem aquesta relació a la xarxa. Un cop decidit que cal incorporar una nova relació al Sistema Temporal, el procés seguit (propagar) és similar a l'exposat a [Allen-83].

La creació d'una nova relació temporal activa l'algorisme de propagació de

relacions de manera que quedin explícites totes les relacions possibles i que no siguin calculables mitjançant el calendari. És necessària la consulta d'una taula de transitivitat que permeti deduir noves relacions, per exemple: essent A, B i C instàncies temporals (totes tres intervals) que no tenen referència directa al calendari i coneixent les relacions A "abans" B i B "solapa" C, podem crear la relació A "abans" C.

En el sistema existeix un esquema singular anomenat XARXAT que conté el nombre de relacions temporals diferents possibles entre elements temporals, el seu codi simbòlic, el seu codi numèric, la taula de relacions inverses i la taula de transitivitat que, donades dues relacions, ens dóna la relació composta. A partir de dues relacions temporals simples és molt freqüent obtenir per composició una relació múltiple, fins i tot és possible obtenir la llista de totes les possibilitats (en aquest cas no es manté cap relació).

La composició de relacions temporals entre intervals és present a molts llocs de la bibliografia especialitzada, en particular a [Allen-83]. Per aquest motiu a la figura 6.7 es mostra una part de la composició de relacions pròpies de COTEM. Les entrades horitzontals corresponen a les possibles relacions entre un interval i un punt. Les entrades verticals corresponen a les possibles relacions entre punts. La composició d'aquests tipus de relacions dóna també relacions entre interval i punt. Quan el resultat de la composició és el conjunt de totes les relacions possibles, ho denotem amb un asterisc.

	PAP	PDP	PIP	PEP	PEPP
IAP	IAP	*	IAP	IAP	IAP, IEP
IDP	*	IDP	IDP	IDP, IEPP	IDP
IEP	IAP	IDP, IEPP, ICPP, ICP	IEP	IAP	IFP
IEPP	ICP, IFP, IEP, IAP	IDP	IEPP	ICPP	IDP
ICPP	ICP, IFP, IEP, IAP	IDP	ICPP	ICP, IFP	IEPP
IFP	IAP	ICP, ICPP, IEPP, IDP	IFP	IEP	ICP, ICPP
ICP	ICP, IFP, IEP, IAP	ICP, ICPP, IEPP, IDP	ICP	ICP, IFP	ICP, ICPP

Figura 6.7

Quan es vol afegir una nova relació temporal entre dos elements temporals, tant si s'ha tret a partir del text com si és resultat de la propagació, es comprova si és consistent amb la relació (múltiple o no) coneguda. Entenem que dues relacions són consistentes quan la seva intersecció no és buida.

Com que la relació entre nodes pot ser múltiple, designem per R_{ab} qualsevol relació bàsica entre els nodes A i B mentre que M_{ab} representa el conjunt de relacions bàsiques presents a la xarxa entre els nodes A i B. Donats M_{ab} i M_{bc} , el càlcul de M_{ac}

es fa aplicant la taula de transitivitat a totes les parelles de relacions formades per un element de M_{ab} i un altre de M_{bc} .

Per facilitar la descripció de l'algorisme, essent R_1 i R_2 relacions bàsiques, considerem que $TR(R_1, R_2)$ representa l'entrada a la taula de transitivitat per aquestes dues relacions. Dessimem per X la instància temporal que acabem de crear i per RD_{xk} la relació, detectada durant l'anàlisi del text, entre X i l'element temporal K . RD_{xk} pot ser una relació múltiple.

Algorisme Actualitzar-Xarxa-Temporal

Inici

si no relacions-detectades-en-el-text-per- X llavors STOP fsi;

{ Tenim X oberta o tancada i relacions-detectades-en-el-text }

si X tancada llavors

 eliminar de relacions-detectades aquelles on K també és tancada;

per cada RD_{xk} de relacions-detectades fer

$AUX_{xk} := RD_{xk} \cap M_{xk}$;

 si RD_{xk} consistent amb M_{xk} llavors

 si $AUX_{xk} \subseteq M_{xk}$ llavors

$M_{xk} := AUX_{xk}$;

 propagar (X, K, AUX)

 fsi

 sino tractament-inconsistències

 fsi

fper

fi

L'acció de propagar té tres paràmetres que defineixen dos nodes i la relació que s'acaba d'incorporar al sistema temporal.

```

Acció Propagar (A, B, NOVA)
Inici
Afegir-a-cua-NOVAab;
mentre cua-no-buida fer
    agafar-NOVAab-de-la-cua;
    per cada C relacionat amb B fer
        si A tancada i C tancada
            llavors passar-a-següent-C
        sino
            NOVAac := Mac  $\cap$  TR (NOVAab, Mbc);
            si NOVAac consistent amb Mac
                llavors si NOVAac  $\subseteq$  Mac llavors
                    Mac := NOVAac;
                    afegir-a-cua-NOVAac
                fsi
            sino tractament-inconsistències
        fsi
    fper;
    per cada C relacionat amb A fer
        si B tancada i C tancada
            llavors passar-a-següent-C
        sino
            NOVAcb := Mcb  $\cap$  TR (Mca, NOVAab)
            si NOVAcb consistent amb Mcb
                llavors si NOVAcb  $\subseteq$  Mcb llavors
                    Mcb := NOVAcb;
                    afegir-a-cua-NOVAcb
                fsi
            sino tractament-inconsistències
        fsi
    fper
fmentre
fi

```

El temps d'execució de l'algorisme d'actualització és més curt que en el cas de mantenir totes les relacions temporals explícites. El motiu és doble: (1) aturada de l'algorisme davant un cert tipus d'instàncies i (2) nombre inferior de relacions a considerar dins l'acció de propagació. En el cas pitjor, totes les instàncies són obertes, no hi ha cap millora. En el cas òptim, totes les instàncies són tancades, les relacions s'obtenen ràpidament via calendari en temps de consulta. El domini de treball és més proper al cas òptim doncs s'utilitza molta informació temporal amb referència directa al calendari.

Les relacions entre instàncies temporals que tenen associats elements del calendari s'obtenen en temps de consulta, moment en el qual s'activa un algorisme de càlcul de relacions que aprofita l'estructura del calendari per deduir-ne ràpidament les relacions que ens interessin.

A continuació descrivim el procediment utilitzat quan es consulta si dues instàncies de la xarxa temporal estan relacionades. La funció definida té dos paràmetres, els nodes objecte de consulta, i retorna la relació entre ells o un fals si no els ha pogut relacionar.

```

Funció Relació (X,Y) retorna resposta
Inici
si Mxy  $\neq \emptyset$  llavors retorna Mxy fsi;
si X tancada i Y tancada llavors retorna càlcul-Mxy-via-calendari fsi;
si X tancada o Y tancada llavors si X tancada llavors
    retorna tercera-instància (X,Y)
sino
    retorna les-inverses-de (tercera-instància (Y,X))
fsi
fsi;
{ Estem en el cas de X i Y obertes i Mxy no existeix }
buscar (X,P);
buscar (Y,Q);
si P i Q existeixen llavors
    càlcul-Mpq-via-calendari;
    Mxq := TR (Mxp,Mpq);
    retorna TR (Mxq,Mqy)
sino retorna fals
fsi
fi

```

```

Funció Tercera-Instància (T,O) retorna resposta
{ El primer paràmetre ha de ser una instància tancada i el segon una oberta }
Inici
  buscar (O,Z);
  si Z existeix llavors
    calcular-Mzt-via-calendari;
    retorna TR(Moz,Mzt)
  sino retorna fals
fsi
fi

```

L'acció buscar(A,B) comprova si existeix una instància tancada B, relacionada amb la instància oberta A, mitjançant la qual es puguin calcular relacions no explícites entre A i altres instàncies. Fem notar que aquesta acció no necessita aplicar els criteris de transitivitat de les relacions temporals. Aixó és degut a com funciona l'algorisme de propagació de relacions.

7. Funcionament de COTEM

La informació continguda en una notícia passa per tres etapes fins a quedar totalment incorporada al sistema. En una primera etapa, es realitza un procés d'anàlisi sintàctica-semàntica de la notícia i s'obté un conjunt d'unitats d'anàlisi que constitueixen una representació del significat de tipus gramàtica de casos. Aquestes unitats són les dades d'entrada per la segona etapa, el processador, en la qual s'extreu la informació relevant en forma d'esquemes similars als continguts a la Base de Coneixement i el Sistema Temporal. La tercera etapa consisteix en la incorporació d'aquests esquemes intermitjos al conjunt del sistema. A la figura 7.1 es presenta l'esquema de funcionament.

La primera i la segona etapa constitueixen el procés d'anàlisi complet que considerem ha de combinar tota la informació de la que disposem: morfològica, sintàctica, semàntica, del context, del domini. La primera etapa consisteix en la combinació de l'anàlisi morfològica, sintàctica i semàntica del text. D'aquesta etapa obtenim una representació del significat de la notícia per *frames* verbals. Com l'objectiu fonamental del treball era l'estudi de la informació temporal, de la primera etapa només s'ha implementat la interpretació semàntica del temps. Si bé el seu desenvolupament complet és factible utilitzant els analitzadors presentats a [Rodríguez-89]. Queden per la segona etapa les qüestions relacionades amb el domini. D'aquesta fase no considerarem el problema de la resolució de les referències, que abordarem en el futur amb l'aplicació dels mecanismes presentats a [DiRoVe-89]. A continuació es descriuen el processador i l'incorporador.

7.1 El processador

Com ja hem mencionat, les dades d'entrada per aquesta fase són les unitats d'anàlisi resultants de la primera etapa: la representació del significat del text (excepte el temps), descomposat en *frames* verbals, similar a una gramàtica de casos.

L'objectiu del procesador és generar a partir d'aquestes unitats d'anàlisi, els corresponents esquemes de la Base de Coneixement i del Sistema Temporal que l'incorporador afegirà com a definitius si són consistents. Els *frames* verbals es corresponen amb esquemes o amb descriptors d'esquemes de la Base de Coneixement, per tant, durant aquesta fase, cal extreure de cada *frame* verbal la informació relevant per tal de crear instàncies d'esquemes o completar algun dels esquemes en curs. Així doncs, el procés d'interpretació utilitza les següents fonts de coneixement:

- El tipus de verbs, de temps verbals i la seva combinació amb el coneixement sobre els tipus de notícies i amb el coneixement sobre el món econòmic.

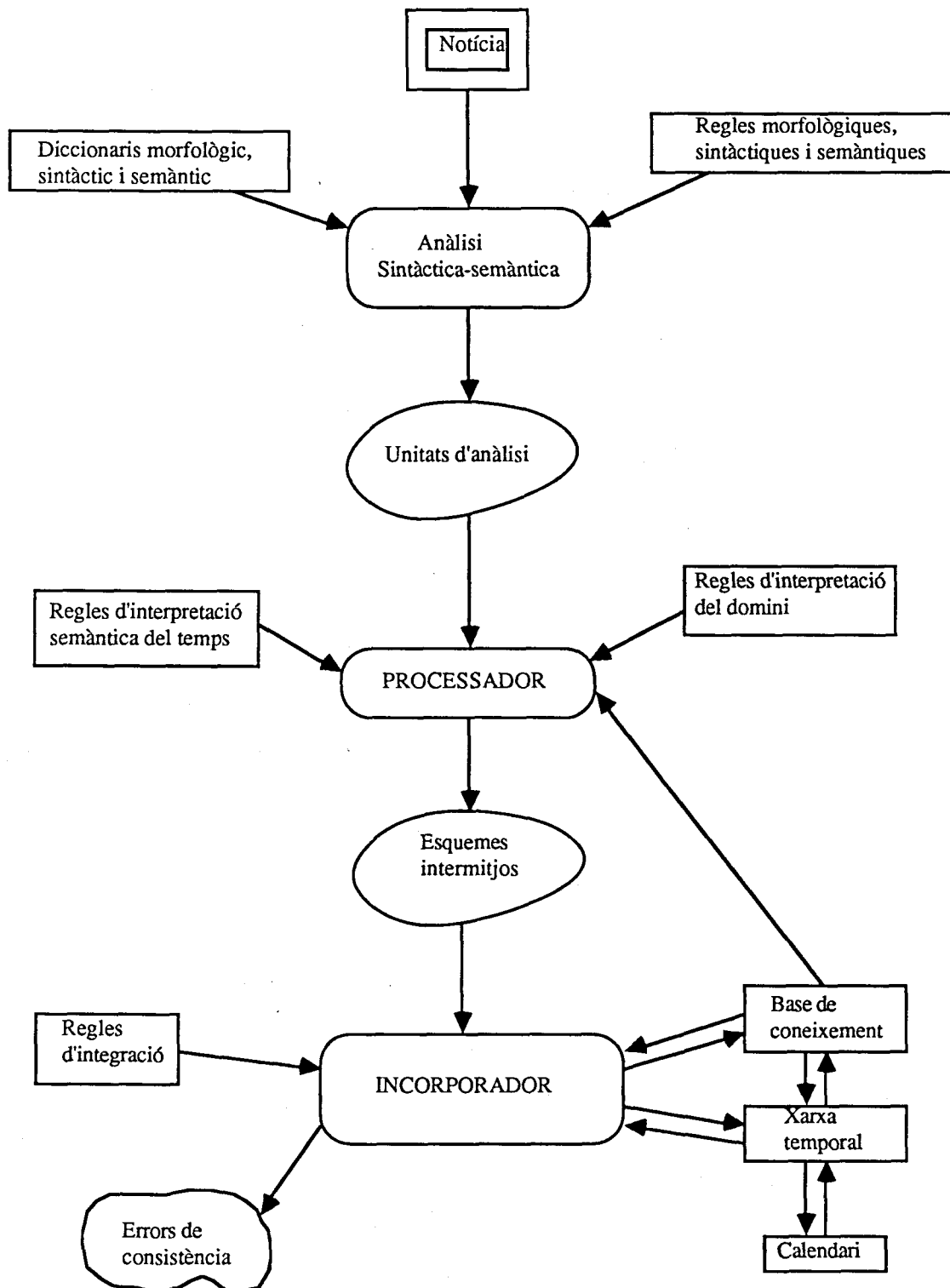


Figura 7.1

- Les relacions que existeixen entre tipus de verbs i tipus de successos econòmics.
- Al mateix temps, el processador ha de:
- Tenir en compte que hi ha verbs que serveixen per crear esquemes, d'altres serveixen només per omplir descriptors i d'altres tenen funcions auxiliars.
 - Considerar el temps verbal com indicador de la funció estructural del verb dins el context de la notícia.
 - Traduir la informació temporal en elements temporals, compaginant tot el coneixement presentat en el capítol 4: les estructures E-R-S dels temps verbals i les expressions temporals.

Dos esquemes auxiliars són el centre d'aquest procés: “resultats-analitzador” i “text-en-curs”. El primer és una instància de seqüència i en el descriptor “en-curso” conté inicialment la llista dels *frames* verbals principals en l'ordre que apareixen en el text (aquells *frames* verbals que són el “tema” d'un altre no apareixen en aquesta llista). L'esquema “text-en-curs” conté, per una banda, una informació inicial referent a la notícia en curs (l'identificador de la notícia, la data de la notícia i el dia de la setmana) i, per una altra banda, emmagatzema les referències a altres esquemes auxiliars: uns que serveixen de control del processador (“notícia” i “altres-notícies”) i altres que serviran de punt de partida a l'incorporador (“incorporar-principal” i “incorporar-secundari”).

El processador extreu un a un els *frames* verbals de “resultats-analitzador” utilitzant els mètodes de tractament de seqüències. Passarà control al *frame* en curs activant, segons quina sigui la situació, un dels dos possibles mètodes associats als *frames* verbals. Abans de passar a la descripció detallada del procés i de cara a facilitar-ne la comprensió, fem una explicació d'aquests mètodes.

En general, cada *frame* verbal té associat dos mètodes: “regles de creació” i “regles de completesa”. Aquests mètodes poden ser particulars de cada *frame* verbal terminal, és a dir, de cada *frame* representant d'un grup de verbs sinònims. Però també poden ser definits de manera més general. Per exemple, els verbs corresponents a la subclasse resum-gestió han sigut dividits en tres grups. Un primer grup constituït pels verbs que presenten el tipus de resum econòmic en el tema (per exemple: *La empresa X ha acumulado unos beneficios de Y millones de pesetas en el pasado año*) i que comparteixen tant “regles de creació” com “regles de completesa”. Un segon grup de verbs el formen aquells que presenten el tipus de resum econòmic a l'experimentador (per exemple: *Los beneficios de la empresa X se elevaron a Y millones de pesetas en el pasado año*) i que també comparteixen entre ells ambdós mètodes. Finalment, el tercer grup consta dels verbs que no són generalitzables pel fet de que el tipus de resum econòmic ve indicat pel propi lexema verbal (per exemple: *La empresa X ha facturado Y*

millones de pesetas en el pasado año) i que per tant necessiten d'un tractament similar als altres verbs de tipus gestió, és a dir, caldrà consultar el descriptor "activa-succes" per saber a quin tipus d'esquema cal referir-se.

Tots els mètodes, dels més generals als més particulars, comparteixen procediments bàsics de l'estil de creació d'elements temporals, combinació d'informació temporal, comprovació de la compatibilitats entre quantitats...

En el cas dels verbs de tipus descripció, només existeix el mètode de completar doncs són verbs que només serveixen per omplir descriptors. També els verbs de tipus gestió-mínima o resultativa posseeixen només el mètode de completar. A l'apartat 3.2 ha sigut presentada la tipologia verbal.

Quan s'activa el mètode "regles de creació" d'una instància d'un *frame* verbal de la subclasse gestió, es crea una instància del succés indicat pel descriptor "activa-succes". En el cas dels resum-gestió pot ser que, en lloc del tipus de succés, hi hagi el mot tema o el mot experimentador, indicant que cal crear una instància de succés del tipus indicat en el corresponent cas verbal. El contingut de la instància de *frame* verbal s'utilitza per omplir descriptors de la instància de succés que s'està creant.

Si s'activa el mètode "regles de creació" per una instància d'un *frame* verbal de tipus administrativa o discursiva, es consulta el valor del cas "tema" que ha de ser també una instància de *frame* verbal i es provoca l'activació del mateix mètode per aquesta segona instància. El més freqüent és que la segona instància ho sigui d'un *frame* verbal de gestió. Però amb les administratives és possible trobar com a "tema" una segona administrativa. Si és aquest el cas, s'activa el mètode "regles de creació" de la instància de *frame* verbal (ara sí de tipus gestió) que estarà com a valor en el descriptor "tema" d'aquesta segona administrativa.

L'activació de "regles de completesa" requereix que s'hagi creat prèviament l'esquema corresponent (un succés, una empresa...). El contingut de la instància de *frame* verbal s'utilitza per completar l'esquema, sigui omplint descriptors buits sigui modificant descriptors amb l'aportació d'informació més precisa.

Quan es creen instàncies de successos econòmics, es considera la possibilitat de que es tracti d'una previsió o d'una realitat. Un succés pot venir marcat com a previsió o es pot deduir que ho és. El cas més evident es produeix quan el succés és presentat utilitzant un operador verbal de previsió (per exemple: *Se prevé una ampliación de capital el próximo año*). En altres casos és el propi verb qui porta l'aspecte de previsió (per exemple: *Los beneficios se estiman en...*). Ara bé, hi ha casos on el concepte de previsió és una qüestió del domini i així ens trobem que el temps verbal futur pot ser un marcador de previsió. Si la notícia parla de resums econòmics en futur, això

és un resultat econòmic previst. En canvi el futur és el temps usual per les notícies que presenten ampliacions de capital ja aprovades i no per aixó li donarem un caire de previsió.

7.1.1 Funcionament del processador.

El primer *frame* verbal correspon al titular. En base a aquest element es decideix de quin tipus estructural és la notícia: text-fet o text-estat (veure l'apartat 3.1).

Un cop decidit el tipus estructural, el processador es guiarà per les expectatives que l'estructura ofereix. El tipus de verb serveix per escollir entre les diferents expectatives que podem tenir en un moment donat.

Recordem que els verbs han sigut tipificats (veure Annex) no només en funció del lexema verbal sinó també d'altres característiques com són el temps verbal, el tema, l'agent o l'experimentador. Per exemple, el verb ampliar es classifica com ampliació-de-capital, resultativa o compra-venda segons el tema sigui capital social, activitats d'empresa o empresa, respectivament.

També cal remarcar que els *frames* verbals que constitueixen la part temàtica de la notícia utilitzen els temps verbals *Pretérito Perfecto*, *Pretérito Indefinido* o *Futuro Perfecto* d'Indicatiu mentre que la part descriptiva utilitza el *Presente* o *Pretérito Perfecto* d'Indicatiu.

Gràficament, el processador tracta els *frames* verbals seguint l'esquema de la figura 7.2.

Les instàncies de *frames* verbals són marcades durant aquesta fase segons la funció que realitzen estructuralment parlant (titular, introducció, descripció-estat...).

Un dels esquemes auxiliars generats contindrà, al final, totes les instàncies de *frames* verbals que representen la notícia organitzades segons l'estructura detectada. Es tracta d'un esquema auxiliar que serveix de control del processador.

L'estudi del titular permet no tan sols decidir el tipus estructural de la notícia sinó també decidir el tipus conceptual de la notícia, si bé a vegades no es pot arribar a la definició més precisa. Les notícies del tipus 1, veure apartat 3.1, són difícilment classificables de manera definitiva consultant només el titular. Se li podrà assignar algun subtipus, però no serà un tipus terminal. Les notícies dels tipus 2 i 3 sí acostumen a ser totalment identificades pel titular.

El processador genera, en funció del tipus conceptual de la notícia, instàncies d'esquemes auxiliars que contindran els esquemes intermitjos que l'incorporador ha de tractar d'afegir a la Base de Coneixement i al Sistema Temporal. Els esquemes intermitjos estaran emmagatzemats com valors en els descriptors de les instàncies d'esquemes auxiliars.

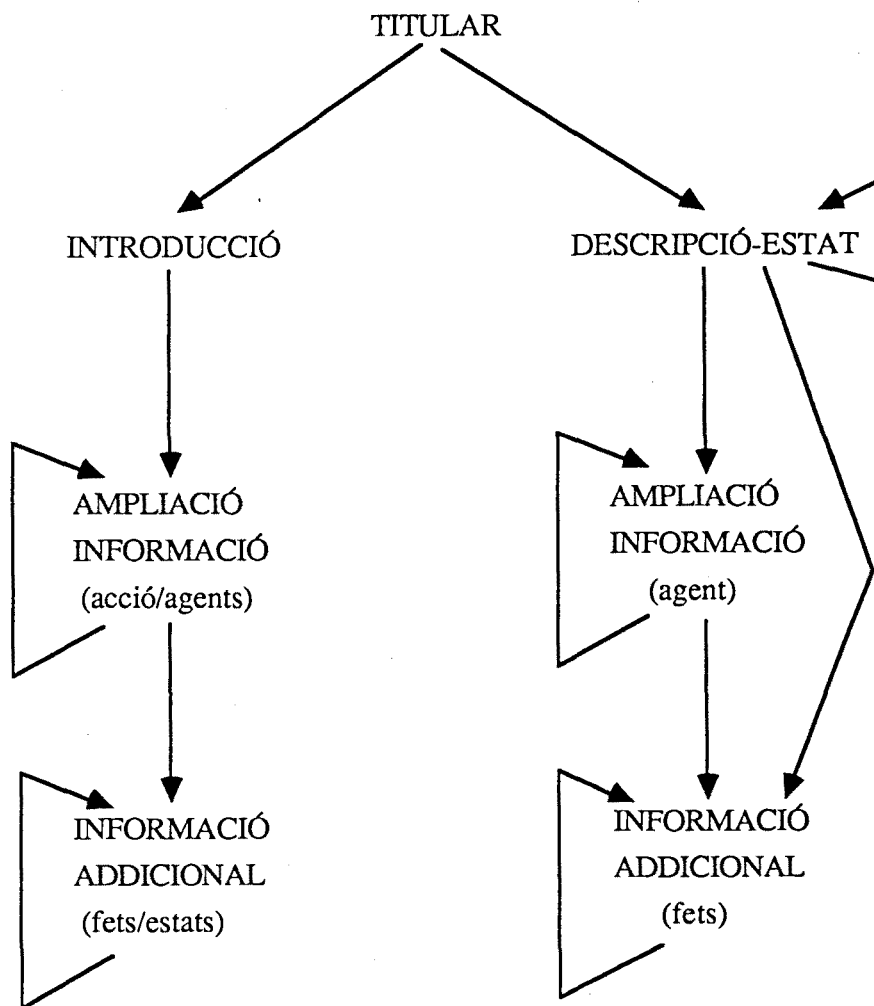


Figura 7.2

A continuació presentem alguns d'aquests esquemes auxiliars, precedits pel més genèric. Aquest esquema conté tres descriptors, els comuns a totes les seves subclasses, dels que serveixen per emmagatzemar esquemes intermitjos.

```

{{ TIPUS-NOTICIA
  ES-UN: ESQ-AUX-NOTICIA
  TIENE-POR-SUBCLASE-A: TN-CANVI-PROPIETAT TN-HIST-EMPRESA TN-ESTATS
  INFORM-TEMPORAL:
  EQUIVALENCIES-TEMPORALS:
  ALTRES-ESQUEMES: }}
  
```

```

{{ TN-CANVI-PROPIETAT
  ES-UN: TIPUS-NOTICIA
  TIENE-POR-SUBCLASE-A:
    TN-TRANSACCIO-COMERCIAL TN-OPERACIO-FINANCERA
  TIPUS-FETS-ECONOMICS-ASSOCIATS:
  AGENTS:
  TEMA:
  OBJECTE:
  NIVELL: NO-TERMINAL}}

{{ TN-HIST-EMPRESA
  ES-UN: TIPUS-NOTICIA
  TIENE-POR-SUBCLASE-A: TN-AMPLIACIO-CAPITAL TN-CREACIO TN-TANCAMENT
  TIPUS-FETS-ECONOMICS-ASSOCIATS:
  AGENTS:
  TEMA:
  NIVELL: NO-TERMINAL}}

{{ TN-ESTATS
  ES-UN: TIPUS-NOTICIA
  AGENT:
  ESTATS:
  PERIODE-PRINCIPAL:
  NIVELL: TERMINAL}}

```

L'estudi del titular condueix a la creació d'una instància d'alguna subclasse de "tipus-noticia". Aquesta instància serà el nou valor pel descriptor "incorporar-principal" de l'esquema "text-en-curs".

Un cop decidit el tipus estructural i el tipus conceptual (amb la precisió permesa), el processador passa el control al mètode "regles de creació" corresponent a la instància de *frame* verbal que constitueix el titular. Aquesta instància correspon a un verb de classe administrativa, discursiva o gestió. Ara es comencen a crear els primers esquemes intermitjos i, en particular, s'omplirà el descriptor "tema" o el descriptor "estats" de la instància de "tipus-noticia" en curs.

Els següents *frames* verbals són tractats d'acord al procés descrit a la figura 7.1 i que podem expressar en forma de regles:

Si (tractat el titular) i (tipus notícia = text-fet)

llavors

Tractar-introducció

Generar expectativa d'ampliació-informació

Durant el tractament de la introducció s'intenta acabar de definir el tipus conceptual de la notícia en els casos que sigui necessari.

Si (tractat el titular) i (tipus notícia = text-estat)

llavors

Tractar-descripció-estat

Generar expectativa de

descripció-estat o ampliació-informació

o informació-addicional

Si (expectativa = ampliació-informació)

llavors

Tractar-ampliació-informació

Generar expectativa de

ampliació-informació o informació-addicional

Aquesta regla només s'aplica un cop en el cas de notícies de tipus text-fet i no s'aplica mai per les notícies de tipus text-estat.

Si (tipus notícia = text-fet) i

(expectativa = (ampliació-informació o informació-addicional)) i

(continua el tema principal)

llavors

Tractar-ampliació-informació

Generar expectativa de

ampliació-informació o informació-addicional

Considerem que un *frame* verbal continua amb el tema principal si és una instància de gestió-mínima o de resultativa o bé fa referència a una gestió (sigui directament per ser una instància de gestió o indirectament per ser una instància d'una administrativa o d'una discursiva de la qual cal consultar el tema per obtenir la gestió) que en el descriptor "activa-succes" té com a valor un dels continguts en el descriptor "tipus-fets-economics-associats" de la instància de "tipus-notícia" en curs.

Si (tipus notícia = text-fet) i
(expectativa = (ampliació-informació o informació-addicional)) i
(NO continua el tema principal)

llavors

Tractar-informació-addicional

Generar expectativa d'informació-addicional

Si (tipus notícia = text-estat) i
(expectativa = (ampliació-informació o informació-addicional)) i
(*frame* verbal en curs és de tipus "descripció")

llavors

Tractar-ampliació-informació

Generar expectativa de

ampliació-informació o informació-addicional

Si (tipus notícia = text-estat) i
(expectativa = (ampliació-informació o informació-addicional)) i
(*frame* verbal en curs NO és de tipus "descripció")

llavors

Tractar-informació-addicional

Generar expectativa d'informació-addicional

Si (expectativa = informació-addicional)

llavors

Tractar-informació-addicional

Generar expectativa d'informació-addicional

Les següents regles només s'apliquen per notícies que són de tipus text-estat.

Si (expectativa =
(descripció-estat o ampliació-informació o informació-addicional)) i
(*frame* verbal en curs és de tipus A)

llavors

Tractar-descripció-estat

Generar expectativa de

descripció-estat o ampliació-informació
o informació-addicional

essent A = “resum-gestió” o “gestió-mínima” de “resum-gestió” o (“administrativa” o “discursiva”) que conté un “resum-gestió”

Si (expectativa =
(descripció-estat o ampliació-informació o informació-addicional)) i
(*frame* verbal en curs és de tipus “descripció”)
llavors
Tractar-ampliació-informació
Generar expectativa de
ampliació-informació o informació-addicional

Si (expectativa =
(descripció-estat o ampliació-informació o informació-addicional)) i
(*frame* verbal en curs NO és de tipus A
i NO és de tipus “descripció”)
llavors
Tractar-informació-addicional
Generar expectativa d’informació-addicional

El processador activa els mètodes “regles de creació” o “regles de completesa” en funció de com decideix considerar cada instància de *frame* verbal. El criteri que segueix és activar “regles de creació” per titulars i informacions-addicionals, activar “regles de completesa” per introduccions, mentre que per ampliacions d’informació i descripcions d’estats activa un mètode o l’altre segons comprovi si es tracta o no del mateix fet (per ampliacions d’informació) o resum (per descripcions d’estats) que l’immediat anterior.

Les informacions addicionals provoquen la creació de noves instàncies de “tipus-notícia” que seran emmagatzemades en el descriptor “incorporar-secundari” de “text-en-curs”.

7.1.2. Processament de la informació temporal

Cada cop que es passa control als mètodes de les instàncies de *frames* verbals, els procediments generals de tractament de la informació temporal seran activats.

Tot succés econòmic (fet o resum) té associat almenys un element temporal. Les ampliacions de capital poden arribar a tenir-ne quatre: el moment d’aplicació, el moment de la proposta, el moment de la decisió i el moment de la supervisió.

El tractament de la informació temporal depèn de l’entorn en el qual s’ha de realitzar: durant la creació d’un succés econòmic o durant la completesa d’un succés econòmic.

Cas de creació d'un succés econòmic.

Es processa tota la informació present a la instància de *frame* verbal per tal de definir l'element temporal corresponent. Si no hi ha cap *quan* a la instància que s'està processant, es crea un element temporal obert del tipus corresponent al succés (punt o interval) i relacionat amb la data de la notícia. Per definir aquesta relació s'utilitza el patró del temps verbal corresponent. Si la instància posseeix algun *quan*, caldrà combinar el coneixement temporal sobre el domini amb tota la informació temporal present a la instància.

El sistema té definits un patró per a cada temps verbal presentat a 4.2.2. Aquests patrons contenen les relacions E-R-S d'una manera genèrica, és a dir, sense especialitzacions per tipus d'element (punt o interval). Cada patró té associat un mètode anomenat "adaptar" que serveix per convertir aquestes relacions genèriques en relacions pròpies del Sistema Temporal: les presentades de manera exhaustiva a l'apartat 6.1.1.

Cas de creació sense quan

Es dedueix el tipus d'element temporal que correspon al succés econòmic (per exemple, un "interval-amb-extrems" per una ampliació de capital, un "punt" per la creació d'una empresa) i se li demana al patró del temps verbal corresponent que "adapti" la relació entre E i S a la situació de que E és del tipus que calgui i S és un punt. El resultat serà, com ja hem dit, que tindrem un element temporal obert relacionat amb la data de la notícia (un element temporal tancat de tipus punt).

Es possible que altres instàncies de *frame* verbal aportin posteriorment més informació que permeti, fins i tot, eliminar aquesta informació temporal tant poc precisa.

Cas de creació amb quan

La situació per aquest cas és diferent. Com ja s'ha comentat a l'apartat 4.2.4, hi ha tipus de *quans* que defineixen E, és a dir, donen directament el moment temporal que ens interessa i altres que proporcionen una referència temporal, defineixen R.

Definir l'element temporal corresponent és immediat en el cas de tenir un *quan* que expressa E. Només cal cridar al mètode "convertir-quan", diferent per cada tipus de *quan*, i ja tenim l'element dessitjat.

Si el *quan* defineix R, tenim que combinar l'expressió temporal i el temps verbal. El procés a seguir ara és més complexe i pot acabar en error. Primer obtenim l'element temporal corresponent al *quan*, per fer això cridem al mètode "convertir-quan". En funció del domini es defineix el tipus d'element temporal associat al succés econòmic. I ara se li demana al patró del temps verbal present en el text que "adapti" les relacions entre R i E segons les restriccions anteriors. Es possible que aquest procés d'adaptació doni error per no ser compatible la definició de R amb la restricció d'E i el temps verbal.

Si es produeix aquesta situació, el processador avisa i es torna a començar considerant que no hi ha *quan*.

La regla genèrica que descriu aquest procés és:

Patró verbal & Valor per R & Restricció per E
→ (relació entre E i R o ERROR)

A continuació presentem diferents exemples d'aplicació d'aquesta regla.

(E durant/comença/acaba/igual R) & valor per R és un punt & E és un punt
→ (E pip R)

En aquest cas fins i tot s'obté un valor concret per E.

(E durant/comença/acaba/igual R) & valor per R és un interval & E és un punt
→ (E ppci/pci/pfi R)

(E durant/comença/acaba/igual R) & valor per R és un punt & E és un interval
→ ERROR

Un exemple d'aquesta situació d'error és la frase *La ampliación de capital se realizará el próximo día 15 de julio*. El sistema espera sempre un interval per aquest tipus de fet econòmic (E és un interval) i es troba amb una expressió temporal que defineix un punt (R és un punt). Donada la seva definició del domini, COTEM no pot entendre aquesta informació temporal.

Cas de completesa d'un succés econòmic.

El tractament de la informació temporal és diferent quan estem sota un procés de completesa. Ara és possible que ens trobem amb la necessitat de combinar la informació temporal present (element temporal nou) en un *frame* verbal amb la informació temporal anterior en el text (element temporal conegut) i que ja està emmagatzemada en el corresponent esquema intermig.

En el cas de que l'element temporal nou s'hagi deduït dels temps verbal sense ajut de cap expressió temporal, aquest element temporal no és té en compte per processos de combinació. Per tenir-lo en compte cal que s'hagi obtingut a partir d'una expressió temporal (un *quan*), combinada o no amb el patró verbal.

Distingirem quatre situacions diferents de combinació d'elements temporals. Remarquem que per definir l'element temporal nou s'ha usat una expressió temporal mentre que de l'element temporal conegut no podem afirmar com s'ha deduït.

- A) L'element temporal conegut s'havia deduït només d'un temps verbal, és a dir, no hi havia cap expressió temporal. En aquests casos es dóna preferència a l'element temporal nou per haver-se deduït d'una expressió temporal.

- A.1) L'element temporal conegut i el nou són inconsistents. Es produeix un missatge d'avís i s'elimina l'element temporal conegut. Queda l'element temporal nou.
- A.2) L'element temporal conegut i el nou són consistents. S'elimina l'element temporal conegut i queda l'element temporal nou.
- B) L'element temporal conegut s'havia obtingut a partir d'una expressió temporal.
 - B.1) L'element temporal conegut i el nou són inconsistents. Si l'element temporal nou és tancat i el conegut és obert, es produeix un missatge d'avís, s'elimina el conegut i ens quedem amb el nou. El motiu és que considerem més probable que la inconsistència vingui d'una expressió poc definida (element temporal obert) i no pas d'una ben definida (element temporal tancat). Per la resta de casos, es produeix un missatge d'avís i es deixa l'element temporal conegut.
 - B.2) L'element temporal conegut i el nou són consistents. Si tots dos són tancats, queda l'element temporal conegut i prou. En els altres casos, el sistema es queda amb tots dos elements temporals, afegint una nova entrada en el descriptor "equivalències-temporals" de la instància de "tipus-notícia" que hi hagi en curs. Si l'element temporal conegut és obert i el nou tancat, es permuten tots dos elements de manera que a l'esquema intermig corresponent queda com a valor el nou element temporal (això simplifica les possibles combinacions amb informació temporal posterior). L'incorporador s'encarregarà d'organitzar totes les possibles equivalències temporals que s'hagin produït.

Durant la fase del processador, per tal de facilitar l'eliminació d'informació quan convingui, tota la informació temporal es crea de manera independent. Això vol dir, per una banda, que al final d'aquesta fase poden existir elements temporals tancats iguals i, per altra, que la propagació de relacions s'ha activat al mínim (només, de manera interna, quan s'han implementat relacions E-S o E-R). L'incorporador haurà de resoldre aquestes situacions.

7.2 L'incorporador

L'objectiu d'aquesta fase és la integració a la Base de Coneixement i al Sistema Temporal de tots els esquemes intermitjos generats durant la fase anterior. Recordem que fins aquí tot el tractament de la notícia s'ha fet situats en el context auxiliar CONTEXT-TREBALL. L'incorporador ha de passar d'aquest context al superior,

CONTEXT-ECONOMIA, tots els esquemes que constitueixen la integració de la notícia al conjunt del sistema.

Per obtenir els esquemes intermitjos l'incorporador ha de consultar l'esquema "text-en-curs". Així obté del descriptor "incorporar-principal" la instància de "tipus-notícia" que correspon al tema central mentre que del descriptor "incorporar-secundari" obté les instàncies de "tipus-notícia" creades durant el tractament de la possible informació addicional. Es fa un tractament individualitzat de cadascuna d'aquestes instàncies amb les comprovacions de consistència pertinents i, si tot ha anat bé, hi ha un procés final per eliminar esquemes sinònims (agents, objectes i successos econòmics) i per unificar els elements temporals tancats iguals.

Tractament de cada esquema intermig

Per cada esquema que l'incorporador troba per primer cop i que representa un agent o un objecte d'un succés econòmic, es comprova si aquest agent ja existia a la Base de Coneixement o no. Si no existeix, és un esquema candidat a ser inclòs si tot va bé. Si existeix, cal comprovar si ambdós esquemes són compatibles. En cas de compatibilitat, la possible nova informació passa a l'esquema conegut i es crea una entrada nova a la relació de parelles d'esquemes sinònims. Si no hi ha compatibilitat, es produeix un missatge d'avís i l'incorporador s'atura.

Per la incorporació dels successos econòmics, es passa control als mètodes associats a cadascun del tipus. Els resums econòmics tenen associat un mètode anomenat "afegir-resum", les ampliacions de capital tenen el mètode "afegir-ampliació"...

Per cada succés econòmic es comprova si es tracta d'una realitat o d'una previsió.

Incorporació d'una previsió

Les previsions són marcades com a tals i, sense més control, quedaran connectades a l'agent (o agents) corresponent(s) via metaconeixement. Així doncs una previsió de beneficis per una empresa anirà al metadescriptor del descriptor "hist-guanyat" d'aquesta empresa. Es decideix que un succés és una previsió bé perquè així ha quedat marcat després de la fase del processador bé perquè les dades de que es disposa, un cop processada tota la notícia, només ens permeten considerar el succés com a previsió (per exemple, una ampliació de capital serà considerada com a previsió si el text no ens ha donat informació sobre l'import de l'ampliació o el percentatge o el nou total de capital social).

Incorporació d'una realitat

Els successos que no són previsions cal comparar-los amb els ja coneguts per veure si es tracta d'un succés nou o no. El criteri per saber si dos successos són el

mateix depèn de cada tipus de succés. Dos resums econòmics del mateix subtipus (beneficis, inversions...) són iguals si corresponen a la mateixa empresa i al mateix període temporal. Els resums de vendes o de pagaments de dividendes requereixen una condició més: igualtat en el “tema-vendes” o en el “per-tipus-accio”, respectivament. Dues ampliacions de capital són la mateixa si corresponen a la mateixa empresa, els seus períodes d'aplicació admeten la igualtat entre elements temporals i hi ha coincidència d'importos de l'ampliació o dels percentatges o dels nous capitals socials.

Quan un mètode d'afegir succés decideix que dos successos són el mateix, tracta de completar l'esquema ja existent amb la informació de l'esquema nou. En cas de repetició d'informació, fa comprovacions de consistència. Per exemple, si tenim l'import d'un resum econòmic tant a l'esquema nou com a l'existent a la Base de Coneixement, es comprova si són quantitats iguals o compatibles. Si això no és compleix, es produeix un error d'incorporació. Si són iguals, ja va bé i, si són compatibles, es fa una estudi per tal de conservar la xifra més precisa. Un altre exemple de combinació d'informació és el cas de dos esquemes referents a la mateixa ampliació de capital i que els seus moments d'aplicació admeten la igualtat, però no són explícitament iguals. En aquest cas s'afegeix explícitament la relació d'igualtat entre tots dos elements temporals, provocant la propagació de relacions temporals corresponent, i les possibles equivalències temporals que tenia l'element temporal de l'esquema nou passen a l'element temporal de l'esquema ja existent a la Base de Coneixement.

Si dos esquemes fan referència al mateix succés i tot el procés de combinació de la informació ha anat bé, aquesta parella d'esquemes passa a incorporar-se a la llista d'esquemes sinònims.

Fase final de l'incorporador

Un cop l'incorporador ha passat per tots els esquemes que calia considerar i tot el procés ha tingut èxit, li queden pendents dues tasques: tractament dels esquemes sinònims i tractament dels elements temporals tancats iguals. Si aquestes dues tasques s'executen sense error, l'incorporador pujarà al CONTEXT-ECONOMIA tots els esquemes nous presents a CONTEXT-TREBALL així com els ja coneguts que han sigut modificats per les noves informacions, quedant la notícia totalment integrada al sistema.

El tractament d'esquemes sinònims consisteix senzillament en eliminar tots els esquemes intermitjos dels quals s'ha trobat la parella a la Base de Coneixement. D'aquesta manera queden només els esquemes que ja es coneixien, però modificats segons la informació aportada pels intermitjos. El tractament dels elements temporals tancats es descriu tot seguit.

7.2.1 Tractament de la informació temporal per part de l'incorporador

Durant la fase d'incorporació hi ha diferents nivells de tractament de la informació temporal: per cada succés econòmic pot ser necessari un procés de combinació d'informació temporal i sempre s'ha de fer el tractament de les possibles equivalències temporals que tingui emmagatzemades la instància de "tipus-notícia" corresponent, per tot el text de la notícia cal fer un tractament final per eliminar elements temporals tancats iguals.

Succés econòmic: Combinació de la informació temporal

A nivell de cada succés econòmic hi ha combinació d'informació temporal només en el cas de que el succés ja sigui conegut pel sistema. Aquesta situació no es dona per tot tipus de succés. Per exemple, com en el cas de resums econòmics s'exigeix igualtat de moments temporals per tal de considerar dos resums iguals, no es produeix la situació de combinació d'elements temporals. En canvi, per les ampliacions de capital es requereix només la possibilitat de que els moments d'aplicació siguin iguals i per tant aquí hi ha un procés de combinació a fer que ja s'ha comentat abans.

Succés econòmic: Tractament d'equivalències temporals

Quan un succés econòmic és vàlid per la seva integració al sistema, cal fer un darrer pas: el tractament, per tot element temporal associat al succés, de les possibles equivalències temporals (parelles d'elements temporals). Aquest tractament consisteix en assignar la relació temporal d'igualtat entre els elements de cada parella, activant la propagació de relacions. Si durant aquest procés es produeixen errors, l'incorporador s'aturarà.

Notícia: Unificació d'elements temporals tancats iguals

Si tots els esquemes intermitjos han sigut tractats i estan a punt per ser integrats al sistema, cal fer un últim tractament d'informació a nivell de tot el text de la notícia. L'incorporador recull totes les instàncies d'element temporal que siguin tancades i tracta d'unificar les que són iguals.

Unificar els elements temporals tancats es fa en dues etapes. La primera etapa unifica, entre ells, els elements temporals tancats iguals presents a CONTEXT-TREBALL. Durant la segona etapa, els elements temporals tancats que han quedat en el context auxiliar són unificats amb els elements temporals tancats del context superior, CONTEXT-ECONOMIA.

Unificar dos elements temporals tancats (A i B) vol dir deixar només un dels dos elements temporals (A), però conservant tota la informació d'ambdós. Hi ha dos tipus d'informació que cal juntar: l'econòmica i la temporal. Per passar la informació

econòmica de l'element B a l'element A, cal copiar les relacions que posseeix B amb els diferents successos econòmics i afegir-les a les que possiblement tingui A. Aquest procés no produeix errors. Passar la informació temporal del B a l'A sí que pot produir errors. El pas d'aquesta informació consisteix en agafar la llista d'elements temporals amb qui està relacionat B, així com les relacions mantingudes, i afegir un a un la relació entre A i cadascun d'aquest elements. El procés de propagació de relacions serà activat i per tant poden aparèixer inconsistències.

Si el procés d'unificació temporal no produeix errors, l'incorporador integra definitivament la notícia a la Base de Coneixement i al Sistema Temporal, donant per finalitzat el tractament de la notícia.

7.3 Reflexions sobre el grau de generalitat

Exposats el processador i l'incorporador, fem ara balanç del nivell de generalitat de cadascun d'ells.

Hem vist que el processador decideix, en base al titular, el tipus estructural de la notícia i, si pot, el tipus conceptual. A partir d'aquí actua en funció de les expectatives estructurals, passant control als *frames* verbals. Si canvia la tipologia de les notícies, el processador haurà de treballar amb diferents expectatives. Si canvia el domini, els mètodes associats als *frames* verbals seran diferents. En qualsevol cas, el funcionament serà el mateix: decidir quina és l'expectativa vàlida en cada moment i passar el control al mètode escaient associat al *frame* verbal en curs.

La implementació actual del processador considera que, donades vàries expectatives, sempre pot decidir quina és la vàlida i només n'hi ha una. Si es permet la possibilitat de tenir diferents expectatives vàlides en un mateix moment, caldrà implementar mecanismes de cerca que explorin l'arbre de possibilitats.

El tractament de la informació temporal és també de caràcter general. Per cada domini caldrà definir la tipologia de les expressions temporals, però el procés de tractament serà com l'exposat a 7.1.2.

Igualment el tractament que fa l'incorporador sobre els elements temporals té un caire general. Tot el que es descriu a 7.2.1 és aplicable a qualsevol domini. Pel que fa a l'incorporació de la informació no temporal, per cada concepte en cada domini s'han de definir els criteris d'integració. Un cop definits, l'incorporador actuarà de manera similar a la descrita anteriorment.

7.4 Exemples de funcionament

En aquest apartat descrivim el processament complet de dues notícies, una d'elles correspon a les de tipus "text-fet" i l'altra a les de tipus "text-estat". Presentem, en

format de Knowledge Craft, els esquemes que constitueixen la representació semàntica de cada text, els esquemes resultants de la fase del processador i els esquemes que deixa com a definitius l'incorporador. Tot això acompanyat de comentaris explicatius.

7.4.1 Una ampliació de capital

Com a exemple de notícia de tipus "text-fet" agafem el text de la figura 7.3 que té, com a part central, informació sobre una ampliació de capital.

El Banco de Santander ampliará su capital en 2.775 millones de pesetas

Madrid

El Banco de Santander realizará una ampliación de su capital social, con cargo a reservas y sin desembolso alguno para los accionistas, por un importe de 2.775,75 millones de pesetas. La decisión de la entidad de proceder a esta ampliación fue adoptada por la junta general del banco, celebrada el pasado 8 de febrero.

Este aumento de capital, con efectos a partir del próximo 12 de abril, se llevará a cabo elevando el valor nominal de todas las acciones en circulación, pasando de las 440 pesetas actuales a 470 pesetas. El capital social del Banco de Santander a 31 de diciembre del pasado año sumaba 40.711 millones de pesetas. Las reservas de la entidad se elevaban en la misma fecha a 70.847 millones de pesetas.

Figura 7.3

Els següents esquemes, FV1 fins FV11, constitueixen la representació interna de la notícia. Els esquemes que apareixen com a valor d'algun descriptor són presentats més endavant.

FV1 representa el titular.

```
{{ FV1
  INSTANCIA-DE: AMPLIAR-1
  ES-UN: INSTANCIA
  TRIPLETA-VERBAL: (T-FUTUR-IMPERFET)
  AGENT: E1
  EN-VALOR: IMP1}}
```

FV2 correspon a la primera frase del text i dóna una mica més d'informació que el titular. Cal remarcar que en el text apareix un operador verbal buit, *realizar*, acompanyant una acció nominalitzada, *ampliación*. En aquests casos l'entrada que reb el sistema és una combinació d'informació de l'estil:

realizar + ampliación → ampliar

passant la informació temporal del verb operador a l'acció.

```
{{ FV2
  INSTANCIA-DE: AMPLIAR-1
  ES-UN: INSTANCIA
  TRIPLETA-VERBAL: (T-FUTUR-IMPERFET)
  AGENT: E1
  EN-VALOR: IMP2
  CARREGADA-A: "reservas"}}
```

FV3, FV4 i FV5 representen la informació sobre la decisió de la ampliació de capital.

```
{{ FV3
  INSTANCIA-DE: ADOPTAR-2
  ES-UN: INSTANCIA
  TEMA: FV4
  TRIPLETA-VERBAL: (T-INDEFINIT)
  AGENT: ORG3
  QUAN-AGENT: EXP-QUAN-1}}
```

```
{{ FV4
  INSTANCIA-DE: DECIDIR
  ES-UN: INSTANCIA
  TEMA: FV5}}
```

```
{{ FV5
  INSTANCIA-DE: AMPLIAR-1
  ES-UN: INSTANCIA}}
```

FV6 i FV7 aporten informació temporal sobre el moment d'aplicació de l'ampliació de capital utilitzant un sinònim d'aquest fet (elevar el valor nominal de les accions). També apareix aquí un operador verbal buit que ha sigut combinat amb l'acció nominalitzada:

llevar a cabo + ampliación → ampliar

{{ FV6
INSTANCIA-DE: AMPLIAR-1
ES-UN: INSTANCIA
TRIPLETA-VERBAL: (T-FUTUR-IMPERFET)
CAS-QUAN: EXP-QUAN-2
COM: FV7}}

{{ FV7
INSTANCIA-DE: ELEVAR-1
ES-UN: INSTANCIA
TRIPLETA-VERBAL: (T-GERUNDI)}}}

FV8 i FV9 donen detalls a nivell bàsic sobre l'ampliació de capital.

{{ FV8
INSTANCIA-DE: AMPLIAR-1
ES-UN: INSTANCIA
TRIPLETA-VERBAL: (T-FUTUR-IMPERFET)
COM: FV9}}

{{ FV9
INSTANCIA-DE: PASAR-DE-A
ES-UN: INSTANCIA
TRIPLETA-VERBAL: (T-GERUNDI)
DES-DE-VALOR: IMP3
FINS-VALOR: IMP4}}

FV10 representa informació addicional. Es tracta d'un resum econòmic: el capital social de l'agent a una data concreta.

{{ FV10
INSTANCIA-DE: ELEVARSE
ES-UN: INSTANCIA
TRIPLETA-VERBAL: (T-IMPERFET)
CAS-QUAN: EXP-QUAN-3
BENEFICIARI: E1
EXPERIMENTADOR: TOTAL-CAPITAL
TEMA: IMP5}}

FV11 també representa informació addicional. Es tracta d'un altre resum econòmic: les reserves (bancàries) de l'agent a una data concreta.

```
{{ FV11
  INSTANCIA-DE: ELEVARSE
  ES-UN: INSTANCIA
  TRIPLETA-VERBAL: (T-IMPERFET)
  CAS-QUAN: EXP-QUAN-3
  BENEFICIARI: E1
  EXPERIMENTADOR: RESERVES
  TEMA: IMP6}}
```

Els esquemes secundaris referenciats per les intàncies de "frame-verbal" són els següents:

```
{{ E1
  INSTANCIA-DE: BANC
  NOM: "Banco de Santander"}}

{{ ORG3
  INSTANCIA-DE: JUNTA
  DE-P-JURIDICA: E1}}

{{ ESPEC1
  INSTANCIA-DE: QUAN-ESPECIFIC
  PUNT-INTER: PUNT
  TIPUS-EXPL: DIA
  VALOR-EXPL: 8
  TIPUS-REF: MES
  VALOR-REF: "febrero"
  POSICIO: PASSAT}}

{{ ESPEC2
  INSTANCIA-DE: QUAN-ESPECIFIC
  PUNT-INTER: PUNT
  TIPUS-EXPL: DIA
  VALOR-EXPL: 12
  TIPUS-REF: MES
  VALOR-REF: "abril"
  POSICIO: FUTUR}}
```

{{ ESPEC3
INSTANCIA-DE: QUAN-ESPECIFIC
PUNT-INTER: PUNT
TIPUS-EXPL: DIA
VALOR-EXPL: 31
TIPUS-REF: MES
VALOR-REF: "diciembre"
POSIC-REF: A-1}}

{{ EXP-QUAN-1
INSTANCIA-DE: QUAN-DE
EXPRESSIO: ESPEC1}}

{{ EXP-QUAN-2
INSTANCIA-DE: QUAN-DES-DE
EXPRESSIO-DES-DE: ESPEC2}}

{{ EXP-QUAN-3
INSTANCIA-DE: QUAN-A-DIRIGIT
EXPRESSIO: ESPEC3}}

{{ IMP1
INSTANCIA-DE: IMPORT
XIFRA: 2775000000
MONEDA: PESETA}}

{{ IMP2
INSTANCIA-DE: IMPORT
XIFRA: 2775750000
MONEDA: PESETA}}

{{ IMP3
INSTANCIA-DE: IMPORT
XIFRA: 440
MONEDA: PESETA}}

{{ IMP4
INSTANCIA-DE: IMPORT
XIFRA: 470
MONEDA: PESETA}}

```
{{ IMP5
  INSTANCIA-DE: IMPORT
  XIFRA: 40711000000
  MONEDA: PESETA}}
```

```
{{ IMP6
  INSTANCIA-DE: IMPORT
  XIFRA: 70847000000
  MONEDA: PESETA}}
```

Per completar la representació del text, manca la presentació dels dos esquemes ja mencionats a la descripció del processador.

L'esquema que conté algunes dades identificatives de la notícia i que servirà per emmagatzemar dades durant tot el tractament de la notícia:

```
{{ TEXT-EN-CURS
  ES-UN: ESQ-AUX-NOTICIA
  IDENTIFICADOR-NOTICIA: "ampliacio-1"
  DATA-NOTICIA: DATA-AMPL-1
  DIA-SETMANA: "miercoles"
  INCORPORAR-PRINCIPAL:
  INCORPORAR-SECUNDARI:
  RELACIO-FINAL-ESQUEMES:
  ESQUEMES-SINONIMS:}}
```

```
{{ DATA-AMPL-1
  INSTANCIA-DE: PUNT
  ES-UN: INSTANCIA
  TIPUS: TANCAT
  VALOR: |d05mar1986|}}
```

```
{{ |d05mar1986|
  INSTANCIA-DE: DIA
  DURANT: |q1mar1986|
  FILL-VIA: DURANT}}
```


L'esquema que indica l'ordre dels *frames* principals que constitueixen la representació de la notícia:

```
{{ RESULTATS-ANALITZADOR
  ES-UN: ESQ-AUX-NOTICIA SECUENCIA
  CLASE-COMPONENTES: FRAME-VERBAL
  EN-CURSO: (FV1 FV2 FV3 FV6 FV8 FV10 FV11)}}
```

El processador tracta un a un els elements de la llista del descriptor "en-curso" d'aquest esquema realitzant tots els treballs descrits a l'apartat 7.1.

Tractament de FV1:

El primer element de la seqüència sempre és el titular de la notícia i així es marca a l'esquema FV1.

```
{{ FV1
  INSTANCIA-DE: EB-TITULAR AMPLIAR-1
  .....  }}
```

El tractament d'aquest *frame* és més complex que els altres doncs, com ja hem comentat, cal decidir el tipus estructural i el tipus conceptual de la notícia. El processador decideix que es tracta d'un "text-fet" doncs AMPLIAR-1 és un verb de tipus gestió, però no del subtipus resum-gestió. Per aquest motiu TEXT-EN-CURS hereta dos descriptors, "notícia" i "altres-notícies", que bàsicament serveixen de control del procesador. En quan al tipus conceptual, consultant el descriptor "activa-succes" d'AMPLIAR-1 es pot deduir que la notícia és del tipus 3.2, és a dir, un subtipus del conjunt de notícies que informen sobre fets de la vida de l'empresa. Per tant es crea una instància de la subclasse corresponent de "tipus-notícia", T10, que anirà emmagatzemant tots els esquemes intermitjos corresponents al tema central de la notícia. Aquesta instància és el nou valor pel descriptor "incorporar-principal" de TEXT-EN-CURS.

Pel que fa al tractament de FV1 com a *frame* verbal qualsevol, s'activa el mètode "regles de creació" associat a AMPLIAR-1 i es crea una instància del fet econòmic "ampliacio-capital", AMPLI00001, que té com "import-ampliacio" IMP1 i com "moment-aplicacio" un element temporal obert de tipus "interval-amb-extrems", IA0001, relacionat amb la data de la notícia.

L'única informació temporal de la que es disposa és el temps verbal que ens situa el fet en relació a S (la data de la notícia). No hi ha cap expressió temporal que ens defineixi E ni tan sols R. En aquest cas el processador guarda la relació E-S. Per coneixement del domini, es restringeix el tipus d'element temporal per E: el moment de

realització d'una ampliació de capital es sempre un interval de temps definit per extrems i que no acostuma a coincidir amb intervals naturals de calendari, per això es crea un "interval-amb-extrems". La regla de combinació aplicada és:

(E després/enllaçat-per S) & E és un interval \rightarrow (E idp/iepp R)

I per tant es crea IA0001 i es modifica DATA-AMPL-1 de la manera següent:

```
{{ IA0001
  INSTANCIA-DE: INTERVAL-AMB-EXTREMS
  ES-UN: INSTANCIA
  TIPUS: OBERT
  RELACIONAT-AMB: (DATA-AMPL-1)
  RELACIONS-MANTINGUDES: (((21 23) T))
  MOMENT-APLICACIO+INV: AMPLI00001}}

{{ DATA-AMPL-1
  INSTANCIA-DE: PUNT
  ES-UN: INSTANCIA
  TIPUS: TANCAT
  VALOR: |d05mar1986|
  RELACIONAT-AMB: (IA0001)
  RELACIONS-MANTINGUDES: (((13 15) T))}}
```

Tractament de FV2:

Tenim un "text-fet" i per tant el processador marca FV2 com introducció.

```
{{ FV2
  INSTANCIA-DE: EB-INTRODUCCIO AMPLIAR-1
  ..... }}
```

Ara s'activa el mètode "regles de completesa" associat a AMPLIAR-1 i la instància creada abans, AMPLI00001, es veu modificada. S'afegeix valor a un descriptor que no en tenia, "carregada-a", i es modifica el valor del descriptor "import-ampliació". En el segon cas, tenim una situació d'informació repetida, però que no és exactament igual: IMP1 i IMP2. La nova informació, IMP2, és més precisa que l'anterior i per tant la substitueix.

Aquest *frame* verbal no aporta nova informació temporal respecte a l'anterior i no hi ha modificacions sobre "moment-aplicació".

Tractament de FV3:

Del tractament de FV2 ha quedat l'expectativa de que a continuació pot venir ampliació de la informació central o informació addicional. Com FV3 té per "tema" a

FV4 i FV4 té per “tema” a FV5 que és una instància d'AMPLIAR-1, el processador dedueix que es tracta d'una ampliació d'informació i així ho marca.

```
{{ FV3
  INSTANCIA-DE: EB-AMPLIACIO-INFORMACIO ADOPTAR-2
  ..... }}
```

S'activa el mètode “regles de completesa” associat a ADOPTAR-2. Com el “tema” és per segona vegada una acció de tipus administrativa, el que succeeix és que les dades passen d'un *frame* a l'altre i s'activa el mètode “regles de completesa” de DECIDIR.

```
{{ FV4
  INSTANCIA-DE: DECIDIR
  ES-UN: INSTANCIA
  TRIPLETA-VERBAL: (T-INDEFINIT)
  TEMA: FV5
  AGENT: ORG3
  EN-LLOC: NIL
  CAS-QUAN: NIL
  QUAN-AGENT: EXP-QUAN-1}}
```

```
{{ FV5
  INSTANCIA-DE: AMPLIAR-1
  ES-UN: INSTANCIA}}
```

Amb el tractament de FV4 podem deduir el moment i l'agent de la decisió del fet indicat per “tema”: FV5 ens porta a AMPLI00001. Per tant “decidida-per” valdrà ORG3. El moment de la decisió s'obté combinant el temps verbal, l'expressió temporal present a “quan-agent” i les restriccions de domini.

Tenim una expressió temporal associada a l'agent, situació raonable per agents del tipus “junta”, “consejo”... Es considera que aquesta expressió defineix R. Per part del domini tenim la restricció de que el moment de decisió (E) dels fets és un punt. Quan hi ha definició per R, el sistema guarda la relació E-R. La regla de combinació aplicada és:

(E igual/durant/comença/acaba R) & (R = D08feb1986) & E és un punt → (E pip R)

En aquest cas ni tan sols es guarda la relació E-R doncs ha quedat E totalment definit. Es crea una instància tancada de tipus “punt”, P0001, que té per valor l'element de calendari |d08feb1986|. Aquesta instància és el valor per “moment-decisio” d'AMPLI00001.

Tractament de FV6:

Del tractament de FV3 ha quedat l'expectativa de que a continuació pot venir ampliació de la informació central o informació addicional. Com FV6 continua parlant de la mateixa acció de gestió, es tracta d'una ampliació d'informació.

```
{{ FV6
  INSTANCIA-DE: EB-AMPLIACIO-INFORMACIO AMPLIAR-1
  .....  }}
```

El valor de "com" és FV7 que resulta ser un *frame* verbal que no diu res. Però FV6 aporta una nova informació molt important: una expressió temporal de les classificades com que defineixen E. Tenim una definició del "moment-aplicacio" de l'ampliació de capital.

Es crea un element temporal obert de tipus "interval-amb-extrems", IA0002, relacionat amb P0002, un element temporal tancat de tipus "punt" (12 d'abril de 1986). La relació temporal és P0002 comença (17) IA0002.

Ara cal combinar IA0001 amb IA0002. Atenent al criteri de prioritat dels elements temporals deduïts usant expressions temporals sobre els deduïts usant només temps verbals, s'elimina IA0001 quedant substituït per IA0002.

Tractament de FV8:

Del tractament de FV6 ha quedat l'expectativa de que a continuació pot venir ampliació de la informació central o informació addicional. També ara es tracta d'una ampliació d'informació.

```
{{ FV8
  INSTANCIA-DE: EB-AMPLIACIO-INFORMACIO AMPLIAR-1
  .....  }}
```

No aporta informació temporal nova, però sí tenim més detalls de l'ampliació. En el descriptor "com" apareix FV9 que resulta ser una instància d'una gestió-mínima associada a ampliacions de capital. Ens permet omplir el descriptor "valor-accio" d'AMPLI0001.

Tractament de FV10:

L'expectativa actual és d'ampliació d'informació o d'informació addicional. FV10 no descriu fets, ens dóna un resum econòmic de l'agent d'AMPLI0001. Tenim ara una informació addicional.

```
{{ FV10
  INSTANCIA-DE: EB-ADDITIONAL ELEVARSE
  .....  }}
```

La implementació actual del processador fa que per cada informació addicional es crei una instància de la subclasse corresponent de “tipus-noticia”. Ara es crea l’esquema T12 que emmagatzema els esquemes relatius a aquest resum econòmic. Aquest T12 es posa com a valor en el descriptor “incorporar-secundari” de TEXT-EN-CURS.

S’activa el mètodes “regles de creació” associat a ELEVARSE i el resultat és la creació d’una instància de la subclasse “total-capital” de “resum-economic”, RESUM00001. A aquesta instància passen totes les dades presents a FV10 i, en particular, la informació temporal és suficient per definir el període associat a aquest resum econòmic.

Tenim una expressió temporal típica de resums econòmics, per tant defineix E. Per obtenir el valor de “moment-total-capital” només cal convertir-la a element temporal. Es crea un element temporal tancat de tipus “punt”, P0003, que té per valor l’element de calendari |d31des1985|.

Tractament de FV11:

Després d’una informació addicional, l’expectativa és de més informacions addicionals com així succeeix.

```
{{ FV11
  INSTANCIA-DE: EB-ADDITIONAL ELEVARSE
  .....  }}
```

També ara es crea una instància de la subclasse corresponent de “tipus-noticia”, T13, que s’afegeix com a valor en el descriptor “incorporar-secundari” de TEXT-EN-CURS.

S’activa el mètode “regles de creació” associat a ELEVARSE i el resultat és la creació d’una instància de la subclasse “reserves” de “resum-economic”, RESUM00002. A aquesta instància passen totes les dades presents a FV11 i, ara també, la informació temporal és suficient per definir el període associat a aquest resum econòmic.

Tenim la mateixa expressió temporal doncs l’analitzador s’haurà encarregat de substituir *en la misma fecha per a 31 de diciembre del pasado año*. Per obtenir el valor de “moment-reserves” només cal convertir-la a element temporal. Es crea un element temporal tancat de tipus “punt”, P0004, que té per valor l’element de calendari |d31des1985|.

Finalitzat el treball del processador, els dos esquemes auxiliars bàsics han quedat de la següent manera:

```
{{ RESULTATS-ANALITZADOR
ES-UN: ESQ-AUX-NOTICIA SECUENCIA
CLASE-COMPONENTES: FRAME-VERBAL
EN-CURSO: NIL
RECORRIDA: (FV11 FV10 FV8 FV6 FV3 FV2 FV1)}}}
```

```
{{ TEXT-EN-CURS
ES-UN: INSTANCIA ESQ-AUX-NOTICIA
INSTANCIA-DE: TEXT-FET
IDENTIFICADOR-NOTICIA: "ampliacio-1"
DATA-NOTICIA: DATA-AMPL-1
DIA-SETMANA: "miercoles"
INCORPORAR-PRINCIPAL: T10
INCORPORAR-SECUNDARI: T13 T12
RELACIO-FINAL-ESQUEMES:
ESQUEMES-SINONIMS:
NOTICIA: T11
ALTRES-NOTICIES: FV11 FV10}}}
```

Els esquemes auxiliars que apareixen com a valors de descriptors de TEXT-EN-CURS són els següents.

La instància de l'esquema auxiliar que serveix de control del processador.

```
{{ T11
INSTANCIA-DE: PRINCIPAL-FET
ES-UN: INSTANCIA
PRIMER-ELEM: FV1
F-SEGON-ELEM: FV2
F-ALTRES: FV8 FV6 FV3}}}
```

La instància de "tipus-noticia" que recull tota la informació central del text.

```
{{ T10
INSTANCIA-DE: TN-AMPLIACIO-CAPITAL
ES-UN: INSTANCIA
TEMA: AMPLI00001
AGENTS: E1
ALTRES-ESQUEMES: IMP4 IMP3 ORG3 IMP2
INFORM-TEMPORAL: (IA0002 P0002) (P0001)}}}
```

Les instàncies de "tipus-noticia" que recullen les informacions addicionals.

```
{{ T12
INSTANCIA-DE: TN-RESUMS
ES-UN: INSTANCIA
ESTATS: RESUM00001
AGENT: E1
ALTRES-ESQUEMES: IMP5
INFORM-TEMPORAL: (P0003)
PERIODE-PRINCIPAL: P0003}}
```

```
{{ T13
INSTANCIA-DE: TN-RESUMS
ES-UN: INSTANCIA
ESTATS: RESUM00002
AGENT: E1
ALTRES-ESQUEMES: IMP6
INFORM-TEMPORAL: (P0004)
PERIODE-PRINCIPAL: P0004}}
```

Els *frames* verbals ja hem vist que quedaven marcats segons la seva funció estructural. Altres esquemes d'entrada al processador no han sigut modificats (ESPEC1, ESPEC2, ESPEC3, EXP-QUAN-1, EXP-QUAN-2, EXP-QUAN-3, E1, ORG3, IMP3, IMP4, IMP5, IMP6). L'esquema IMP1 ha sigut eliminat en ser substituït per IMP2, al qual s'ha intentat passar-li possible informació continguda a IMP1 i per tant ha quedat així:

```
{{ IMP2
INSTANCIA-DE: IMPORT
XIFRA: 2775750000
MONEDA: PESETA
XIFRA-EQUIV: NIL
MONEDA-EQUIV: NIL}}
```

Els esquemes intermitjos que s'han creat durant aquesta fase representen els successos econòmics i la informació temporal present en el text. Aquests esquemes són relacionats en les pàgines següents.

```
{{ AMPLI00001
  INSTANCIA-DE: AMPLIACIO-CAPITAL
  ES-UN: INSTANCIA
  INTERESSAT: E1
  CARREGADA-A: "reservas"
  PAS-ACCIO: NIL
  REAL-PREVISIO: REAL
  IMPORT-AMPLIACIO: IMP2
  MOMENT-APLICACIO: IA0002
  DECIDIDA-PER: ORG3
  MOMENT-DECISIO: P0001
  VALOR-ACCIO: (IMP3 IMP4 0)}}}
```

```
{{ RESUM00001
  INSTANCIA-DE: TOTAL-CAPITAL
  ES-UN: INSTANCIA
  DE-QUI: E1
  REAL-PREVISIO: REAL
  QUANTITAT-RESUM: IMP5
  MOMENT-TOTAL-CAPITAL: P0003}}}
```

```
{{ RESUM00002
  INSTANCIA-DE: RESERVES
  ES-UN: INSTANCIA
  DE-QUI: E1
  REAL-PREVISIO: REAL
  QUANTITAT-RESUM: IMP6
  MOMENT-RESERVES: P0004}}}
```

Nous elements temporals:

```
{{ P0001
  INSTANCIA-DE: PUNT
  ES-UN: INSTANCIA
  ARGUMENTOS-SLOT+INV:
  TIPUS: TANCAT
  VALOR: |d08feb1986|
  MOMENT-DECISIO+INV: AMPLI00001}}}
```



```

{{ P0002
  INSTANCIA-DE: PUNT
  ES-UN: INSTANCIA
  ARGUMENTOS-SLOT+INV:
  TIPUS: TANCAT
  VALOR: |d12abr1986|
  RELACIONAT-AMB: (IA0002)
  RELACIONS-MANTINGUDES: (((17) T))}}

{{ P0003
  INSTANCIA-DE: PUNT
  ES-UN: INSTANCIA
  ARGUMENTOS-SLOT+INV:
  TIPUS: TANCAT
  VALOR: |d31des1985|
  MOMENT-TOTAL-CAPITAL+INV: RESUM00001}}

{{ P0004
  INSTANCIA-DE: PUNT
  ES-UN: INSTANCIA
  ARGUMENTOS-SLOT+INV:
  TIPUS: TANCAT
  VALOR: |d31des1985|
  MOMENT-RESERVES+INV: RESUM00002}}

{{ IA0002
  INSTANCIA-DE: INTERVAL-AMB-EXTREMS
  ES-UN: INSTANCIA
  TIPUS: OBERT
  RELACIONAT-AMB: (P0002)
  RELACIONS-MANTINGUDES: (((24) T))
  MOMENT-APLICACIO+INV: AMPLI00001}}

```

Element temporal ja existent:

```
{{ DATA-AMPL-1
  INSTANCIA-DE: PUNT
  ES-UN: INSTANCIA
  TIPUS: TANCAT
  VALOR: |d05mar1986|
  RELACIONAT-AMB: NIL
  RELACIONS-MANTINGUDES: NIL}}
```

Finalment presentem els elements de calendari més significatius que s'han creat, però cal recordar que la creació d'un element de calendari pot comportar la creació de tot un arbre d'elements per sobre d'ell fins arribar a l'any.

```
{{ |d08feb1986|
  INSTANCIA-DE: DIA
  ES-UN: INSTANCIA
  DURANT: |q1feb1986|
  ABANS: |d05mar1986|
  FILL-VIA: DURANT
  GERMA-DRET-VIA: ABANS
  VALOR+INV: P0001}}
```

```
{{ |d12abr1986|
  INSTANCIA-DE: DIA
  ES-UN: INSTANCIA
  DURANT: |q1abr1986|
  DESPRES: |d05mar1986|
  FILL-VIA: DURANT
  GERMA-ESQUERRE-VIA: DESPRES
  VALOR+INV: P0002}}
```

```
{{ |d31des1985|
  INSTANCIA-DE: DIA
  ES-UN: INSTANCIA
  ACABA: |q2des1985|
  FILL-VIA: ACABA
  VALOR+INV: P0004 P0003}}
```

Aquest element ja existia i ha tingut alguna modificació pel que fa a les seves relacions amb els altres elements de calendari que s'han creat.

```
{{ |d05mar1986|
INSTANCIA-DE: DIA
DURANT: |q1mar1986|
DESPRES: |d08feb1986|
ABANS: |d12abr1986|
FILL-VIA: DURANT
GERMA-ESQUERRE-VIA: DESPRES
GERMA-DRET-VIA: ABANS}}
```

Ara es posa en marxa l'incorporador que haurà de comprovar si la informació és nova o no i, si no ho és, fer controls de consistència. També ha d'unificar els elements temporals tancats iguals. Resulta ser una notícia amb informació totalment nova pel sistema.

Els successos econòmics queden lligats a l'agent:

```
{{ E1
INSTANCIA-DE: BANC
NOM: "Banco de Santander"
HIST-AMPLIADA: AMPLIO0001
HIST-RESERVES: RESUM00002
HIST-CAPITAL: RESUM00001}}

{{ AMPLIO0001
INSTANCIA-DE: AMPLIACIO-CAPITAL
ES-UN: INSTANCIA
INTERESSAT: E1
CARREGADA-A: "reservas"
PAS-ACCIO: NIL
REAL-PREVISIO: REAL
IMPORT-AMPLIACIO: IMP2
MOMENT-APLICACIO: IA0002
DECIDIDA-PER: ORG3
MOMENT-DECISIO: P0001
VALOR-ACCIO: (IMP3 IMP4 0)}}}
```

```
{{ RESUM00001
  INSTANCIA-DE: TOTAL-CAPITAL
  ES-UN: INSTANCIA
  DE-QUI: E1
  REAL-PREVISIO: REAL
  QUANTITAT-RESUM: IMP5
  MOMENT-TOTAL-CAPITAL: P0004}}
```

```
{{ RESUM00002
  INSTANCIA-DE: RESERVES
  ES-UN: INSTANCIA
  DE-QUI: E1
  REAL-PREVISIO: REAL
  QUANTITAT-RESUM: IMP6
  MOMENT-RESERVES: P0004}}
```

Els elements temporals tancats iguals, P0003 i P0004, queden unificats en P0004 i per tant s'ha modificat RESUM00002.

```
{{ P0004
  INSTANCIA-DE: PUNT
  ES-UN: INSTANCIA
  ARGUMENTOS-SLOT+INV:
  TIPUS: TANCAT
  VALOR: |d31des1985|
  MOMENT-RESERVES+INV: RESUM00002
  MOMENT-TOTAL-CAPITAL+INV: RESUM00001}}
```

També son incorporats a la Base de Coneixement els esquemes ORG3, IMP2, IMP3, IMP4, IMP5 i IMP6 tal com han quedat després del processador.

Els nous elements temporals del Sistema Temporal són P0001, P0002 i IA0002 tal com s'han creat durant la fase anterior, més el P0004 tal com ha quedat després de la unificació.

Els elements de calendari també passen al Sistema Temporal tal com s'han creat durant la fase anterior, excepte l'element relacionat amb P0003 i P0004 que queda així:

```
{{ |d31des1985|
  INSTANCIA-DE: DIA
  ES-UN: INSTANCIA
  ACABA: |q2des1985|
  FILL-VIA: ACABA
  VALOR+INV: P0004}}
```

7.4.2 Uns resultats econòmics

A la figura 7.4 presentem una notícia exemple de les de tipus "text-estat". El text informa sobre alguns resultats econòmics d'un Banc (beneficis i *cash-flow*) i aporta una previsió sobre pagaments de dividendes. El tractament d'aquesta notícia s'ha fet just després de tractar l'anterior i, per aquest motiu, els elements de calendari ja apareixen lligats amb els elements creats abans.

El SGB obtuvo en 1985 unos beneficios de 692 millones de francos suizos

1-3-86

Madrid

El Schweizerische Bankgesellschaft (SBG), la mayor entidad bancaria de Suiza, obtuvo el pasado ejercicio unos beneficios netos de 692 millones de francos suizos (unos 51.200 millones de pesetas), lo que representa un incremento del 19% con respecto a 1984.

El *cash flow* (beneficio más amortizaciones) registró un aumento del 21% (231 millones de francos) sobre 1984, alcanzando un total de 1.348 millones. En la junta general, que tendrá lugar el próximo 1 de abril, el consejo propondrá la distribución de un dividendo de 120 francos por acción al portador y de 24 francos por acción nominativa.

SOLSA EXTRANJERA

Figura 7.4

Els següents esquemes, FV1 fins FV9, són les dades d'entrada pel processador.
FV1 representa el titular.

```
{{ FV1
  INSTANCIA-DE: OBTENER-1
  ES-UN: INSTANCIA
  TRIPLETA-VERBAL: (T-INDEFINIT)
  CAS-QUAN: EXP-QUAN-1
  EXPERIMENTADOR: E2
  TEMA: BENEFICIS
  FINS-VALOR: IMP7}}
```

FV2 i FV3 representen el primer paràgraf, on s'amplia el titular i es compara el resultat econòmic amb el de l'any anterior.

```
{{ FV2
  INSTANCIA-DE: OBTENER-1
  ES-UN: INSTANCIA
  TRIPLETA-VERBAL: (T-INDEFINIT)
  CAS-QUAN: EXP-QUAN-2
  EXPERIMENTADOR: E2
  TEMA: BENEFICIS
  FINS-VALOR: IMP8}}
```

```
{{ FV3
  INSTANCIA-DE: REPRESENTAR-2
  ES-UN: INSTANCIA
  TRIPLETA-VERBAL: (T-PRESENT)
  TEMA: FV3BIS}}
```

```
{{ FV3BIS
  INSTANCIA-DE: INCREMENTO
  ES-UN: INSTANCIA
  TRIPLETA-VERBAL: (T-PRESENT)
  TEMA: P1
  CAS-QUAN: EXP-QUAN-3}}
```

FV4 i FV5 representen la informació sobre una altre resum econòmic que també es compara amb l'any anterior.

```
{{ FV4
  INSTANCIA-DE: AUMENTAR-1-2
  ES-UN: INSTANCIA
  TRIPLETA-VERBAL: (T-INDEFINIT)
  EXPERIMENTADOR: CASH-FLOW
  CAS-QUAN: EXP-QUAN-4
  TEMA: P2}}
```

```
{{ FV5
  INSTANCIA-DE: ALCANZAR-1-2
  ES-UN: INSTANCIA
  TRIPLETA-VERBAL: (T-INDEFINIT)
  EXPERIMENTADOR: CASH-FLOW
  TEMA: IMP9}}
```

FV6, FV7, FV8 i FV9 representen la proposta d'un tercer tipus de resultat econòmic: repartiment de dividendes en funció del tipus d'acció.

```
{{ FV6
  INSTANCIA-DE: PROPONER
  ES-UN: INSTANCIA
  TEMA: FV7
  TRIPLETA-VERBAL: (T-FUTUR-IMPERFET)
  AGENT: ORG2
  BENEFICIARI: ORG1
  QUAN-BENEFICIARI: EXP-QUAN-5}}
```

```
{{ FV7
  INSTANCIA-DE: REPARTIR
  ES-UN: INSTANCIA
  FINS-VALOR: IMP14
  PER-TIPUS-ACCIO: PORTADOR}}
```

{{ FV8
INSTANCIA-DE: PROPONER
ES-UN: INSTANCIA
TEMA: FV9
TRIPLETA-VERBAL: (T-FUTUR-IMPERFET)
AGENT: ORG2
BENEFICIARI: ORG1
QUAN-BENEFICIARI: EXP-QUAN-5}}

{{ FV9
INSTANCIA-DE: REPARTIR
ES-UN: INSTANCIA
FINS-VALOR: IMP15
PER-TIPUS-ACCIO: NOMINATIVA}}

Els esquemes secundaris referenciats pels esquemes anteriors són:

{{ E2
INSTANCIA-DE: BANC
NACIONALITAT: SUISA
NOM: "Schweizerische Banckgesellschaft"}}

{{ ORG1
INSTANCIA-DE: JUNTA
DE-P-JURIDICA: E2}}

{{ ORG2
INSTANCIA-DE: CON-DIRECCIO
DE-P-JURIDICA: E2}}

{{ ESPEC1
INSTANCIA-DE: QUAN-ESPECIFIC
PUNT-INTER: INTERVAL
TIPUS-EXPL: ANY
VALOR-EXPL: 1985}}

{{ ESPEC2
INSTANCIA-DE: QUAN-ESPECIFIC
PUNT-INTER: INTERVAL
TIPUS-EXPL: ANY
POSICIO: PASSAT}}


```

{{ ESPEC3
  INSTANCIA-DE: QUAN-ESPECIFIC
  PUNT-INTER: INTERVAL
  TIPUS-EXPL: ANY
  VALOR-EXPL: 1984}}

{{ ESPEC4
  INSTANCIA-DE: QUAN-ESPECIFIC
  PUNT-INTER: INTERVAL
  TIPUS-EXPL: ANY
  VALOR-EXPL: 1984}}

{{ ESPEC5
  INSTANCIA-DE: QUAN-ESPECIFIC
  PUNT-INTER: PUNT
  TIPUS-EXPL: DIA
  VALOR-EXPL: 1
  POSICIO: FUTUR
  TIPUS-REF: MES
  VALOR-REF: "abril"}}

{{ EXP-QUAN-1
  INSTANCIA-DE: QUAN-EN-INTERVAL
  EXP-INTERVAL: ESPEC1}}

{{ EXP-QUAN-2
  INSTANCIA-DE: QUAN-0
  EXPRESSIO: ESPEC2}}

{{ EXP-QUAN-3
  INSTANCIA-DE: QUAN-GERMA
  PERIODE-ANY: ESPEC3}}

{{ EXP-QUAN-4
  INSTANCIA-DE: QUAN-GERMA
  PERIODE-ANY: ESPEC4}}

{{ EXP-QUAN-5
  INSTANCIA-DE: QUAN-0
  EXPRESSIO: ESPEC5}}

```

```

{{ IMP7
  INSTANCIA-DE: IMPORT
  XIFRA: 692000000
  MONEDA: FRANC-SUIS}}

{{ IMP8
  INSTANCIA-DE: IMPORT
  XIFRA: 692000000
  MONEDA: FRANC-SUIS
  XIFRA-EQUIV: 51200000000
  MONEDA-EQUIV: PESETA}}

{{ IMP9
  INSTANCIA-DE: IMPORT
  XIFRA: 1348000000
  MONEDA: FRANC-SUIS}}

{{ IMP14
  INSTANCIA-DE: IMPORT
  XIFRA: 120
  MONEDA: FRANC-SUIS}}

{{ IMP15
  INSTANCIA-DE: IMPORT
  XIFRA: 24
  MONEDA: FRANC-SUIS}}

{{ P1
  INSTANCIA-DE: PERCENTATGE
  TANT-PER-CENT: 19}}

{{ P2
  INSTANCIA-DE: PERCENTATGE
  TANT-PER-CENT: 21
  VALOR-ABSOLUT: 231000000}}

```

Els esquemes auxiliars pel processador contenen la següent informació:

```

{{ RESULTATS-ANALITZADOR
  ES-UN: ESQ-AUX-NOTICIA SECUENCIA
  CLASE-COMPONENTES: FRAME-VERBAL
  EN-CURSO: (FV1 FV2 FV3 FV4 FV5 FV6 FV8)}}

```

```

{{ TEXT-EN-CURS
  ES-UN: ESQ-AUX-NOTICIA
  IDENTIFICADOR-NOTICIA: "resum-economic-48"
  DATA-NOTICIA: DATA-RESUM-48
  DIA-SETMANA: "sabado"
  INCORPORAR-PRINCIPAL:
  INCORPORAR-SECUNDARI:
  RELACIO-FINAL-ESQUEMES:
  ESQUEMES-SINONIMS:}}

{{ DATA-RESUM-48
  INSTANCIA-DE: PUNT
  ES-UN: INSTANCIA
  TIPUS: TANCAT
  VALOR: |d01mar1986|}}

{{ |d01mar1986|
  INSTANCIA-DE: DIA
  COMENCA: |q1mar1986|
  ABANS: |d05mar1986|
  DESPRES: |d08feb1986|
  FILL-VIA: COMENCA
  GERMA-DRET-VIA: ABANS
  GERMA-ESQUERRE-VIA: DESPRES}}

```

Descrivim a continuació el treball realitzat pel processador: el tractament de cadascun dels elements de la llista del descriptor "en-curso" de RESULTATS-ANALITZADOR. Fem una explicació més breu tenint en compte que molts detalls ja s'han comentat a l'exemple anterior.

Tractament de FV1:

Essent el primer element de la seqüència, cal marcar FV1 com a titular.

```

{{ FV1
  INSTANCIA-DE: EB-TITULAR OBTENER-1
  ..... }}

```

El processador decideix que es tracta d'un "text-estat" doncs OBTENER-1 és un verb de tipus resum-gestió. L'esquema TEXT-EN-CURS hereta els descriptors "noticia" i "altres-fets" que serviran de control del processador. En quan al tipus conceptual no hi ha dubtes: la notícia és de tipus 2, és a dir, pertany al conjunt de notícies

que avaluen l'activitat de les empreses. Es crea T14, instància de "tipus-noticia" que emmagatzemarà els esquemes intermitjos.

El tractament de FV1 com a *frame* verbal qualsevol activa el mètode "regles de creació" associat a OBTENER-1. Es crea RESUM00003 que és una instància del subtipus de "resum-economic" indicat per "tema", en aquest cas "beneficis". La quantitat, IMP7, passa al descriptor corresponent.

El titular presenta una expressió temporal que es considera que defineix E quan acompanya un verb de tipus resum-gestió. Com és aquest el cas, a partir d'ella es crea un element temporal tancat, IS0001, que correspon al moment temporal associat a aquest resum econòmic.

En el cas de notícies de tipus "text-estat" és molt freqüent presentar diferents resultats econòmics corresponents a un mateix període de temps sense mencionar per cadascun d'ells el moment temporal. Per això en aquest tipus de notícies mantenim el primer element temporal accessible per tota la notícia. El IS0001 és el valor assignat al descriptor "període-principal" de T14. Quan un resultat econòmic no té indicació de període, es pren el contingut d'aquest descriptor.

Tractament de FV2:

Tenim un "text-estat" i l'expectativa creada al tractar el titular és de trobar una descripció d'estat que acostuma a ser l'ampliació del titular, però no sempre ho és.

```
{{ FV2
  INSTANCIA-DE: EB-DESCRIPCIO-ESTAT OBTENER-1
  ..... }}
```

En aquest cas FV2 sí és una ampliació del titular. Però totes les dades del resum econòmic són conegudes excepte l'equivalència dels beneficis entre francs i pessetes. El resultat d'aplicar el mètode "regles de completesa" a FV2 és tan sols la modificació d'IMP7 que es veu ampliat amb les dades d'IMP8.

```
{{ IMP7
  INSTANCIA-DE: IMPORT
  XIFRA: 692000000
  MONEDA: FRANC-SUIS
  XIFRA-EQUIV: 51200000000
  MONEDA-EQUIV: PESETA}}
```

Tractament de FV3:

Al tractar FV2 s'ha creat l'expectativa de trobar una nova descripció d'estat o ampliació d'informació o informació addicional. Resulta que REPRESENTAR-2 és una

gestió-mínima de resum-gestió, per tant continuem amb la descripció d'estat.

{{ FV3

INSTANCIA-DE: EB-DESCRIPCIO-ESTAT REPRESENTAR-2

..... }}

Aquest verb és molt típic en aquest tipus de notícies. Es tracta sempre amb el mètode "regles de completesa". Veiem que com a tema té un operador-resum-gestió, FV3BIS, que presenta una comparació del mateix resum econòmic en períodes diferents. Es crea un element temporal tancat, IS0002, que correspon al període anterior. Aquest element i el percentatge present a FV3BIS passen als descriptors escaients del RESUM00003.

Tractament de FV4:

Continuem amb l'expectativa de trobar una nova descripció d'estat o ampliació d'informació o informació addicional. Tenim un verb de tipus resum-gestió, per tant és tracta d'una nova descripció d'estat.

{{ FV4

INSTANCIA-DE: EB-DESCRIPCIO-ESTAT AUMENTAR-1-2

..... }}

Com no es continua amb el mateix resum econòmic, s'activa el mètode "regles de creació" i es crea una nova instància del subtipus de "resum-economic", RESUM00004, indicat ara per "experimentador".

Apareix una expressió temporal que no dona directament el període del resum econòmic, dona el període anterior. Per tant el moment associat a RESUM00004 s'extreu de l'esquema T14 (del descriptor "període-principal"). Es crea un element temporal tancat, IS0003, per indicar el període anterior amb el qual es compara.

Tractament de FV5:

Després de tractar FV4, l'expectativa és de trobar una nova descripció d'estat o ampliació d'informació o informació addicional. Tenim un verb de tipus resum-gestió, per tant es tracta d'una descripció d'estat.

{{ FV5

INSTANCIA-DE: EB-DESCRIPCIO-ESTAT ALCANZAR-1-2

..... }}

És una continuació del resum anterior i per tant s'activa el mètode "regles de completesa". Ens aporta l'import, IMP9, corresponent a RESUM00004.

Tractament de FV6:

Segueix l'expectativa de trobar una nova descripció d'estat o ampliació d'informació o informació addicional. Tenim un verb de tipus administrativa que té com a tema un resum-gestió, per tant es una descripció d'estat.

```
{{ FV6
  INSTANCIA-DE: EB-DESCRIPCIO-ESTAT PROPONER
  ..... }}
```

Es tracta d'un resum econòmic nou i per tant s'activa el mètode de "regles de creació" associat a PROPONER. Tractant FV6 s'obté les dades respecte al moment de la proposta de FV7. En particular es crea un element temporal tancat de tipus "punt", P0007, que correspon al moment de la proposta. La situació de càlcul temporal és similar a la presentada en el tractament de FV4 de l'exemple anterior (7.4.1).

Algunes dades passen a FV7 i s'activa "regles de creació" per REPARTIR. Observar que FV7 ha quedat marcat com a previsió.

```
{{ FV7
  INSTANCIA-DE: REPARTIR
  ES-UN: INSTANCIA
  FINS-VALOR: IMP14
  PER-TIPUS-ACCIO: PORTADOR
  PREVISIO: T
  TRIPLETA-VERBAL: ((T-INDEFINIT))}}
```

Del tractament de FV7 surt una nova instància de "resum-economic" que és RESUM00005. Com no hi ha indicació del moment temporal, s'agafa IS0001. Cal indicar que, mentre el pagament de dividendes és en futur, el període en el qual s'ha obtingut els beneficis i al qual volem associar els dividendes ja ha passat. Això fa que també ara s'apliqui el criteri de que tots els resums presents a una notícia corresponen a un mateix període, si no diuen res en contra, per tal d'obtenir el moment temporal d'aquest repartiment de dividendes.

Tractament de FV8:

El processament de FV8 i FV9 és paral·lel al de FV6 i FV7 doncs provenen de la descomposició d'una conjunció. Es marca FV8:

```
{{ FV8
  INSTANCIA-DE: EB-DESCRIPCIO-ESTAT PROPONER
  ..... }}
```

Es crea una instància de "resum-economic" nova, RESUM00006, i un nou element temporal tancat de tipus "punt", P0008, per representar el moment de la proposta.

En aquesta notícia no apareixen informacions addicionals. El processador ha acabat la seva feina i els esquemes auxiliars bàsics han quedat així:

```
{{ RESULTATS-ANALITZADOR
  ES-UN: ESQ-AUX-NOTICIA SECUENCIA
  CLASE-COMPONENTES: FRAME-VERBAL
  EN-CURSO: NIL
  RECORRIDA: (FV8 FV6 FV5 FV4 FV3 FV2 FV1)}}}
```

```
{{ TEXT-EN-CURS
  ES-UN: INSTANCIA ESQ-AUX-NOTICIA
  INSTANCIA-DE: TEXT-ESTAT
  IDENTIFICADOR-NOTICIA: "resum-economic-48"
  DATA-NOTICIA: DATA-RESUM-48
  DIA-SETMANA: "sabado"
  INCORPORAR-PRINCIPAL: T14
  INCORPORAR-SECUNDARI:
  RELACIO-FINAL-ESQUEMES:
  ESQUEMES-SINONIMS:
  NOTICIA: T15}}}
```

Els esquemes auxiliars T14 i T15 són:

```
{{ T14
  INSTANCIA-DE: TN-RESUMS
  ES-UN: INSTANCIA
  ESTATS: RESUM00006 RESUM00005 RESUM00004 RESUM00003
  AGENT: E2
  ALTRES-ESQUEMES: ORG1 IMP15 ORG1 IMP14 IMP9 P2 P1 IMP7
  INFORM-TEMPORAL: (P0008) (P0007) (IS0003) (IS0002) (IS0001)
  PERIODE-PRINCIPAL: IS0001}}}
```

```
{{ T15
  INSTANCIA-DE: PRINCIPAL-ESTAT
  ES-UN: INSTANCIA
  PRIMER-ELEM: FV1
  E-SEGON-ELEM: FV8 FV6 FV5 FV4 FV3 FV2}}}
```

El marcatge dels *frames* verbals ja s'ha vist durant el procés. També s'ha

presentat la modificació d'IMP7 per IMP8. Els altres esquemes d'entrada no han sigut modificats (E2, ORG1, ORG2, ESPEC1, ESPEC2, ESPEC3, ESPEC4, ESPEC5, EXP-QUAN-1, EXP-QUAN-2, EXP-QUAN-3, EXP-QUAN-4, EXP-QUAN-5, IMP9, IMP14, IMP15, P1, P2). Els esquemes intermitjos creats per representar els resultats econòmics i la informació temporal són els següents:

```
{{ RESUM00003
  INSTANCIA-DE: BENEFICIS
  ES-UN: INSTANCIA
  DE-QUI: E2
  QUANTITAT-RESUM: IMP7
  REAL-PREVISIO: REAL
  MOMENT-BENEFICIS: IS0001
  TPC-RESPECTE-ANTERIOR: P1
  MOMENT-ANTERIOR: IS0002}}

{{ RESUM00004
  INSTANCIA-DE: CASH-FLOW
  ES-UN: INSTANCIA
  DE-QUI: E2
  REAL-PREVISIO: REAL
  TPC-RESPECTE-ANTERIOR: P2
  MOMENT-ANTERIOR: IS0003
  MOMENT-CASH-FLOW: IS0001
  QUANTITAT-RESUM: IMP9}}

{{ RESUM00005
  INSTANCIA-DE: PAGAMENT-DIVIDENDS
  ES-UN: INSTANCIA
  DE-QUI: E2
  QUANTITAT-RESUM: IMP14
  REAL-PREVISIO: PREVIST
  PER-TIPUS-ACCIO: PORTADOR
  MOMENT-PAGAMENT-DIVIDENDS: IS0001
  PROPOSADA-A: ORG1
  MOMENT-PROPOSTA: P0007}}
```



```

{{ RESUM00006
  INSTANCIA-DE: PAGAMENT-DIVIDENDS
  ES-UN: INSTANCIA
  DE-QUI: E2
  QUANTITAT-RESUM: IMP15
  REAL-PREVISIO: PREVIST
  PER-TIPUS-ACCIO: NOMINATIVA
  MOMENT-PAGAMENT-DIVIDENDS: IS0001
  PROPOSADA-A: ORG1
  MOMENT-PROPOSTA: P0008}}

{{ P0007
  INSTANCIA-DE: PUNT
  ES-UN: INSTANCIA
  ARGUMENTOS-SLOT+INV:
  TIPUS: TANCAT
  VALOR: |d01abr1986|
  MOMENT-PROPOSTA+INV: RESUM00005}}

{{ P0008
  INSTANCIA-DE: PUNT
  ES-UN: INSTANCIA
  ARGUMENTOS-SLOT+INV:
  TIPUS: TANCAT
  VALOR: |d01abr1986|
  MOMENT-PROPOSTA+INV: RESUM00006}}

{{ IS0001
  INSTANCIA-DE: INTERVAL-SENSE-EXTREMS
  ES-UN: INSTANCIA
  ARGUMENTOS-SLOT+INV:
  TIPUS: TANCAT
  VALOR: |any1985|
  MOMENT-BENEFICIS+INV: RESUM00003
  MOMENT-CASH-FLOW+INV: RESUM00004
  MOMENT-PAGAMENT-DIVIDENDS+INV: RESUM00006 RESUM00005}}

```

```
{{ IS0002
  INSTANCIA-DE: INTERVAL-SENSE-EXTREMS
  ES-UN: INSTANCIA
  ARGUMENTOS-SLOT+INV:
  TIPUS: TANCAT
  VALOR: |any1984|
  MOMENT-ANTERIOR+INV: RESUM00003}}
```

```
{{ IS0003
  INSTANCIA-DE: INTERVAL-SENSE-EXTREMS
  ES-UN: INSTANCIA
  ARGUMENTOS-SLOT+INV:
  TIPUS: TANCAT
  VALOR: |any1984|
  MOMENT-ANTERIOR+INV: RESUM00004}}
```

L'element temporal ja existent per definir la data de la notícia no ha sigut modificat.

```
{{ DATA-RESUM-48
  INSTANCIA-DE: PUNT
  ES-UN: INSTANCIA
  TIPUS: TANCAT
  VALOR: |d01mar1986|}}
```

Els principals elements de calendari creats durant aquesta fase són els següents:

```
{{ |d01mar1986|
  INSTANCIA-DE: DIA
  COMENCA: |q1mar1986|
  ABANS: |d05mar1986|
  DESPRES: |d08feb1986|
  FILL-VIA: COMENCA
  GERMA-DRET-VIA: ABANS
  GERMA-ESQUERRE-VIA: DESPRES}}
```

{{ |d01abr1986|
INSTANCIA-DE: DIA
ES-UN: INSTANCIA
COMENCA: |q1abr1986|
ABANS: |d12abr1986|
DESPRES: |d05mar1986|
FILL-VIA: COMENCA
GERMA-DRET-VIA: ABANS
GERMA-ESQUERRE-VIA: DESPRES
VALOR+INV: P0008 P0007}}

{{ |any1985|
MIEMBRO-DE: CALENDARI
ENLLACAT-PER: |any1984|
INSTANCIA-DE: ANY
ENLLACA: |any1986|
ES-UN: INSTANCIA
ACABAT-PER: |s21985| |s21985|
COMENCAT-PER: |s11985|
GERMA-DRET-VIA: ENLLACA
GERMA-ESQUERRE-VIA: ENLLACAT-PER
PARE-DE: |s11985| |s21985|
VALOR+INV: IS0001}}

{{ |any1984|
MIEMBRO-DE: CALENDARI
ENLLACAT-PER: |any1983|
INSTANCIA-DE: ANY
ENLLACA: |any1985|
ES-UN: INSTANCIA
GERMA-DRET-VIA: ENLLACA
GERMA-ESQUERRE-VIA: ENLLACAT-PER
VALOR+INV: IS0003 IS0002}}

La posta en marxa de l'incorporador deixa com a definitius els següents esquemes.
Tots els resums econòmics queden lligats a l'agent:

```
{{ E2
INSTANCIA-DE: BANC
NACIONALITAT: SUISA
NOM: "Schweizerische Banckgesellschaft"
HIST-DIVIDENDS:
    [INSTANCE: HIST-DIVIDENDS]
    [SCHEMA: E2]
    [SLOT: HIST-DIVIDENDS]
    [PREVISIONS: RESUM00005 RESUM00006]
HIST-CASH-FLOW: RESUM00004
HIST-GUANYAT: RESUM00003}}
```

Fem notar que les previsions del pagament de dividendes han quedat en el metadescriptor i no directament en el descriptor "hist-dividends".

```
{{ RESUM00003
INSTANCIA-DE: BENEFICIS
ES-UN: INSTANCIA
DE-QUI: E2
QUANTITAT-RESUM: IMP7
REAL-PREVISIO: REAL
MOMENT-BENEFICIS: IS0001
TPC-RESPECTE-ANTERIOR: P1
MOMENT-ANTERIOR: IS0003}}
```

```
{{ RESUM00004
INSTANCIA-DE: CASH-FLOW
ES-UN: INSTANCIA
DE-QUI: E2
REAL-PREVISIO: REAL
TPC-RESPECTE-ANTERIOR: P2
MOMENT-ANTERIOR: IS0003
MOMENT-CASH-FLOW: IS0001
QUANTITAT-RESUM: IMP9}}
```

```
{{ RESUM00005
INSTANCIA-DE: PAGAMENT-DIVIDENDS
ES-UN: INSTANCIA
DE-QUI: E2
QUANTITAT-RESUM: IMP14
REAL-PREVISIO: PREVIST
PER-TIPUS-ACCIO: PORTADOR
MOMENT-PAGAMENT-DIVIDENDS: IS0001
PROPOSADA-A: ORG1
MOMENT-PROPOSTA: P0008}}
```

```
{{ RESUM00006
INSTANCIA-DE: PAGAMENT-DIVIDENDS
ES-UN: INSTANCIA
DE-QUI: E2
QUANTITAT-RESUM: IMP15
REAL-PREVISIO: PREVIST
PER-TIPUS-ACCIO: NOMINATIVA
MOMENT-PAGAMENT-DIVIDENDS: IS0001
PROPOSADA-A: ORG1
MOMENT-PROPOSTA: P0008}}
```

A la Base de Coneixement també s'incorporen ORG1, IMP7, IMP9, IMP14, IMP15, P1 i P2 tal com han quedat després de la fase del processador.

En el procés d'unificació d'elements temporals s'eliminen P0007 i IS0002 quedant com nous elements del Sistema Temporal P0008, IS0001 i IS0003

```
{{ P0008
INSTANCIA-DE: PUNT
ES-UN: INSTANCIA
ARGUMENTOS-SLOT+INV:
TIPUS: TANCAT
VALOR: |d01abr1986|
MOMENT-PROPOSTA+INV: RESUM00005 RESUM00006}}
```

```

{{ IS0001
  INSTANCIA-DE: INTERVAL-SENSE-EXTREMS
  ES-UN: INSTANCIA
  ARGUMENTOS-SLOT+INV:
  TIPUS: TANCAT
  VALOR: |any1985|
  MOMENT-BENEFICIS+INV: RESUM00003
  MOMENT-CASH-FLOW+INV: RESUM00004
  MOMENT-PAGAMENT-DIVIDENDS+INV: RESUM00006 RESUM00005}}

```

```

{{ IS0003
  INSTANCIA-DE: INTERVAL-SENSE-EXTREMS
  ES-UN: INSTANCIA
  ARGUMENTOS-SLOT+INV:
  TIPUS: TANCAT
  VALOR: |any1984|
  MOMENT-ANTERIOR+INV: RESUM00003 RESUM00004}}

```

Els elements de calendari també passen al Sistema Temporal tal com han sortit del processador, excepte els afectats per la unificació d'elements temporals tancats:

```

{{ |d01abr1986|
  INSTANCIA-DE: DIA
  ES-UN: INSTANCIA
  COMENCA: |q1abr1986|
  ABANS: |d12abr1986|
  DESPRES: |d05mar1986|
  FILL-VIA: COMENCA
  GERMA-DRET-VIA: ABANS
  GERMA-ESQUERRE-VIA: DESPRES
  VALOR+INV: P0008}}

```

```

{{ |any1984|
  MIEMBRO-DE: CALENDARI
  ENLLACAT-PER: |any1983|
  INSTANCIA-DE: ANY
  ENLLACA: |any1985|
  ES-UN: INSTANCIA
  GERMA-DRET-VIA: ENLLACA
  GERMA-ESQUERRE-VIA: ENLLACAT-PER
  VALOR+INV: IS0003}}

```


8. Conclusions

El treball descrit en aquesta memòria ha consistit en l'estudi de la problemàtica del tractament de textos narratius, tenint la informació temporal com a objectiu principal: com s'expressa en el text, com s'interpreta, com intervé en el procés de comprensió del text i com es representa en un sistema de coneixement.

Ha sigut necessari definir un domini d'aplicació: les notícies de diari. Treballant sobre aquest domini s'han desenvolupat eines de caràcter general, tant per anàlisi del text com a nivell de mètodes de representació i raonament.

Estudiant el domini s'han definit taxonomies de notícies i taxonomies de verbs. Creiem que la metodologia aplicada és vàlida tant si es canvia de domini com si s'amplia el domini actual.

El punt de partida ha sigut l'estudi dels textos a tractar, tant en forma com en contingut. Això ha permès establir una tipologia de les notícies a nivell estructural i una altra a nivell conceptual. Al mateix temps ha quedat caracteritzat el coneixement sobre el domini, necessari per la total comprensió del text. Aquest estudi és igualment necessari en qualsevol domini on es vulgui realitzar una comprensió automàtica de textos.

La base de coneixement i la tipologia conceptual sempre estaran en funció del domini, com és obvi, però es pot pensar en una tipologia estructural més o menys general, si més no, per les notícies. La modificació o ampliació de les taxonomies no comporta un canvi en el funcionament del sistema. Simplement en el moment de prendre decisions o generar expectatives haurà de tenir en compte l'existència de més alternatives o de tipologies noves i, per tant, haurà de tenir altres criteris de selecció.

La tipologia verbal s'ha fet des d'un punt de vista semàntic i, per tant, depèn del domini. L'ampliació del domini significarà la revisió d'aquesta tipologia. El canvi de domini provocarà, amb tota seguretat, una nova tipologia malgrat que es mantinguin les classes més generals. Per exemple, els verbs discursius no són particulars de les notícies sobre economia, els podem trobar a qualsevol tipus de notícies i, fins i tot, a qualsevol text narratiu.

De tota manera, el sistema tractarà els verbs de la mateixa forma en qualsevol domini. Essent que la representació semàntica d'aquests indica si serveixen per crear instàncies o per modificar instàncies (omplir descriptors), només variarà, situats en cada verb, què és el que es crea o el que es modifica.

Utilitzant un formalisme basat en *frames*, s'ha implementat una Base de Coneixement en la qual es representen els elements del domini (fets o estats econòmics, objectes financers, tipologia conceptual de les notícies...), el coneixement lingüístic

(categorització verbal, casos semàntics, patrons temporals dels temps verbals...) i el coneixement general, amb la inclusió d'un model del temps. Les instàncies d'elements temporals i les relacions entre elles es creen i mantenen apart en el Sistema Temporal.

L'expressió lingüística del temps en castellà s'ha estudiat a dos nivells. A nivell de frase, proposem una classificació de les expressions temporals, els patrons E-R-S per temps verbals en castellà i la combinació de les expressions temporals amb els patrons verbals. A nivell de discurs, les notícies són classificades segons la seva estructura narrativa i hem definit relacions entre aquests tipus estructurals, els lexemes verbals i els temps verbals.

En quan al temps verbals remarquem que l'assignació de les estructures E-R-S a cadascun d'ells s'ha fet en funció del castellà en general i no particularitzant sobre el domini. El que sí s'ha fet és simplificar-ne alguna si d'aquesta manera s'afavoria l'eficiència del sistema. S'han implementat els temps verbals més freqüents en el domini, però els mecanismes de representació i d'interpretació estan oberts a qualsevol nova incorporació.

En quan a les expressions temporals és evident que hem fet una tipificació d'aquelles que apareixen en el domini, però hem tractat de distingir entre les que tenen un caràcter general, vàlides per a qualsevol domini, les que són en certa mida generals dins del domini i les més particulars. Al mateix temps, un canvi de domini comportarà una nova tipologia; però, fent l'estudi en la línia del presentat aquí, els mecanismes d'interpretació i de combinació seran totalment aplicables.

El Sistema Temporal dissenyat és independent del domini i consta d'una xarxa d'elements temporals de diferents tipus (punt, interval) connectada amb un calendari. Els elements temporals són classificats en tancats i oberts segons si tenen o no, respectivament, relació amb elements de calendari. La xarxa manté els elements lligats mitjançant relacions temporals especialitzades. També imposa restriccions en funció de si són oberts o tancats. Aixó fa que l'algorisme de propagació de relacions temporals sigui més eficient.

El prototipus implementat tracta les notícies a partir d'una representació semàntica del text. Les notícies són interpretades (processador) i la seva informació és incorporada (incorporador) a tot el sistema de representació.

El processador decideix, en base al titular, el tipus estructural de la notícia i, si pot, el tipus conceptual. A partir d'aquí actua en funció de les expectatives estructurals, passant control als *frames* verbals. Si canvia la tipologia de les notícies, el processador haurà de treballar amb diferents expectatives. Si canvia el domini, els mètodes associats als *frames* verbals seran diferents. En qualsevol cas, el funcionament serà el mateix:

decidir quina és l'expectativa vàlida en cada en moment i passar el control al mètode escaient associat al *frame* verbal en curs.

El tractament de la informació temporal és també de caràcter general. Per cada domini caldrà definir la tipologia de les expressions temporals, però el procés de tractament no varia.

Igualment el tractament que fa l'incorporador sobre els elements temporals té un caire general. Tot el seu funcionament és aplicable a qualsevol domini. Pel que fa a l'incorporació de la informació no temporal, per cada concepte en cada domini s'han de definir els criteris d'integració. Un cop definits l'incorporador actuarà de manera similar a la descrita anteriorment.

Línies futures

La recerca desenvolupada deixa moltes opcions de continuïtat, des d'ampliacions de qüestions puntuals allà on hem arribat a resultats parcials fins ampliacions del sistema i altres aplicacions.

Queden oberts els següents temes:

- Tractament de les expressions temporals que fan referència a altres successos presents en el text o a la Base de Coneixement.
- Estudi i representació de les cadenes temporals.
- Incorporació del concepte de distància temporal.
- Estudi de l'aspecte verbal en profunditat.
- Disseny d'un sistema de consulta acoplat a COTEM.

Pel que fa al tractament del temps, també és força interessant l'estudi comparatiu de l'aplicació de la nostra proposta a altres tipus de text. Igualment, seria profitós comparar amb les propostes basades en teories de la narrativa com la d'en H. Kamp.

Annex: Verbs del domini i tipus associat

A continuació presentem un subconjunt representatiu dels verbs del corpus de notícies econòmiques.

Per cada verb indiquem el tipus associat i algunes peculiaritats, sigui descriptors que el caracteritzen, sigui restriccions d'ús, ...

El signe \equiv indica sinònim-de.

Elements que permeten determinar el tipus d'una acció:

- 1) El tipus d'agent
- 2) El tipus de tema
- 3) El tipus d'experimentador
- 2) El temps verbal

Els tipus són indicats per un comentari o per un concepte de la Base de Coneixement.

ACOGER Descripció-empresa

ACOGERSE-A Administrativa

ACORDAR Administrativa

ACUMULAR Resum-gestió

ADOPTAR-1 Descripció-empresa

Adoptar nom una empresa. Apareix a fusions o creacions.

ADOPTAR-2 Administrativa

Descriptor QUAN obligatori, amb el valor de 'data concreta' (un dia).
Present directament o indirectament.

ADQUIRIR Compra-venda

Agent: Persona

AFECTAR Descripció-gestió

ALCANZAR-1 Resum-gestió

ALCANZAR-2 Administrativa

AMPLIAR-1 Ampliació-de-capital

Tema: capital social

AMPLIAR-2 Resultativa

Tema: activitats d'empresa

AMPLIAR-3 Compra-venda

Tema: empresa

Instrument: compra

ANTICIPAR ≡ ANUNCIAR

ANUNCIAR Discursiva

APROBAR Administrativa

ASCENDER ≡ ELEVARSE

ASEGURAR-1 Discursiva

Tema: Una oració subordinada substantiva

ASEGURAR-2 Resultativa

Experimentador: persona

Tema: nominalització (“puesto de trabajo”, “continuidad”, ...)

AUMENTAR-1 Resum-gestió

Regeix un QUAN-GERMA (pot ser per defecte)

AUMENTAR-2 ≡ AMPLIAR-1

AUMENTO ≡ INCREMENTO

CELEBRAR Administrativa

CERRAR Tancament

COMERCIALIZAR Activitat-empresa

COMPRAR-1 Compra-venda (Si és una nominalització, en singular)

COMPRAR-2 Resum-gestió

COMPROMETERSE Operador-verbal-qualificador

comprometerse + ‘a’ + acció infinitiu → acció en futur

CONFIRMAR-1 Administrativa ≡ RATIFICAR

Agent: òrgan empresa

CONFIRMAR-2 Discursiva

Agent: d’altres

CONSEGUIR-1 ≡ OBTENER-1

CONSEGUIR-2 ≡ ALCANZAR-2

CONSIDERAR (ús impersonal) ≡ TRATAR-DE (ambdues opcions)

CONSOLIDAR Resultativa

CONSTITUIR-1 Creació

CONSTITUIR-2 ≡ SER-2

CONSTITUIR-3 (estar constituido por) ≡ FORMAR

CONTAR-CON Descripció-empresa

CONTINUAR-EN Resultativa

CONVERTIR-EN Resultativa

Agent: Gestió nominalitzada

Tema: empresa

En-que: resultat. Descriptor específic d'aquesta acció. És un SP regit per 'en'.

CORRESPONDER-A Gestió Mínima (de Resum-gestió)

CREAR Creació

CRECER-1 \equiv INCREMENTARSE

CRECER-2 Gestió-mínima (d'Ampliació-de-capital)

CRECIMIENTO \equiv INCREMENTO

DAR-A-CONOCER Discursiva

DAR-EMPLEO \equiv EMPLEAR-A

DECIDIR Administrativa

DECLARARSE-EN-BANCARROTA Tancament

DESCENDER Resum-gestió

DESEMBOLSAR Gestió-mínima (de Compra-venda i d'Ampliació-de-capital)

DESINVERTIR \equiv VENDER-1

DISTRIBUIR \equiv REPARTIR

DUPLICAR Resum-gestió

Regeix un QUAN-GERMA encara que sigui per defecte

EFECTUAR \equiv HACER

ELEVAR-1 Ampliació-de-capital

Tema: "valor nominal de las acciones"

Opcional "de ... a ..."

ELEVAR-2 / ELEVARSE Resum-gestió

Si es parla del capital social, trobarem que dins de textos d'ampliació es presenta el verb en futur per indicar el nou capital social i sempre que es descriu un resum-gestió es presenta en imperfet.

EMITIR Gestió-mínima (d'Ampliació-de-capital)

Tema: acciones, títulos, ...

EMPLEAR-A Descripció-empresa

ENTRAR-EN Operador-verbal-qualificador

entrar-en + acció durativa nominalitzada \rightarrow marca l'inici de l'acció

EQUIVALER \equiv REPRESENTAR-2
ESPERAR-1 Administrativa
ESPERAR-2 Discursiva
ESTAR-PENDIENTE-DE \equiv ESPERAR-1
ESTIMAR (ús impersonal) Resum-gestió (amb caire de previsió)
ESTUDIAR Discursiva
EXPANDIRSE [EXPANSION de (empresa) en (sector)] Resultativa
EXPORTAR Activitat-empresa
FABRICAR Activitat-empresa
FACTURAR Resum-gestió
FIRMAR Administrativa
FORMAR (estar formado por) Descripció-empresa
FUSIONAR / FUSION Fusió
HACER Operador-verbal-buit
HACER-PÚBLICO \equiv ANUNCIAR
INCORPORAR Compra-venda
INCORPORARSE Fusió / Integració
INCREMENTAR / INCREMENTARSE \equiv AUMENTAR-1
INCREMENTO (sempre nominalitzat) Operador-resum-gestió
de (---%) {sobre/respecto} QUAN
INFORMAR-1 Discursiva
Agent: Agent ben definit
INFORMAR-2 Destructiva
Agent: Agent indefinit
INSCRIBIR \equiv ADOPTAR-1
INSTALAR Gestió-mínima (de Creació)
INTEGRARSE Integració
INTRODUCIRSE Resultativa
INVERTIR-1 Gestió-mínima (de Compra i de Comanda)
(Si és una nominalització, en singular)
INVERTIR-2 Resum-gestió

LLEGAR-A-1 ≡ ELEVARSE
LLEGAR-A-2 ≡ ALCANZAR-2
LLEVAR-A-CABO ≡ HACER
MEJORAR Resultativa
MODERNIZAR Resultativa
OBTENER-1 Resum-gestió
OBTENER-2 Administrativa
OCUPAR Descripció-empresa
OCUPAR-A ≡ EMPLEAR-A
OFRECER Compra-venda
Tema: compra
Freqüentment s'usa nominalitzat
OPERAR-EN Descripció-empresa
OSTENTAR ≡ POSEER
PARECER Destructiva
PARTICIPAR Gestió-mínima (d'Ampliació-de-capital)
PASAR DE ... A ... Gestió-mínima (d'Ampliació-de-capital)
PEDIR-1 Comanda
Agent: Persona
PEDIR-2 Administrativa ("Pedir autorización para ampliación")
PERDER Resum-gestió
PERTENECER Descripció-gestió
PONER-A-LA-VENTA Compra-venda
PONER-DE-RELIEVE Resultativa
POSEER Descripció-empresa
PREVER Operador-verbal-qualificador
prever + {acció nominalitzada/acció infinitiu} → acció prevista
PRODUCIR-1 Resum-gestió
PRODUCIR-2 Activitat-empresa
PROPONER Administrativa

PROPONERSE Operador-verbal-qualificador
 proponerse + acció nominalitzada → acció en futur

QUEDAR-FIJADO Gestió-mínima (d'Ampliació-de-capital)
 Tema: Capital-social

QUERER Operador-verbal-qualificador
 proponerse + acció infinitiu → acció en futur

RATIFICAR Administrativa

REALIZAR ≡ HACER

RECIBIR-1 Gestió-mínima (de Comanda)

RECIBIR-2 Administrativa

REESTRUCTURAR Gestió-mínima (de Tancament)

REGISTRAR ≡ TENER-1

REMATAR [una operació] Administrativa

RENEGOCIAR Gestió-mínima (de Tancament)

REPARTIR Resum-gestió
 Tema: "Dividendos"

REPRESENTAR-1 Resultativa

REPRESENTAR-2 Gestió-mínima (de Resum-gestió)
 Sempre porta un operador-resum-gestió

RESPONDER-A Descripció-gestió

REVELAR-1 Discursiva

REVELAR-2 Resultativa

RUMOREAR [SE VENÍA RUMOREANDO] Destructiva

SEÑALAR Discursiva

SER-1 Descripció-empresa

SER-2 Descripció-gestió

SER-FRUTO/RESULTADO-DE Descripció-gestió

SER-SUPERIOR-A xxxx EN yyyy Gestió-mínima (de Resum-gestió)

SITUARSE-1 Descripció-empresa

SITUARSE-2 ≡ ELEVARSE

SOLICITAR Administrativa
 Beneficiari: Organisme

SUMAR-1 Descripció-gestió
SUMAR-2 ≡ ELEVARSE
SUPERAR Gestió-mínima (de Resum-gestió)
SUPONER ≡ REPRESENTAR (ambdues opcions)
SUSCRIBIR-1 Administrativa
SUSCRIBIR-2 Gestió-mínima (d'Ampliació-de-capital)
Tema: acciones
TENER-1 Resum-gestió
TENER-2 Descripció-empresa
TENER-INTENCIÓN-DE Operador-verbal-qualificador
tener-intención-de + acció infinitiu → acció en futur
TENER-PREVISTO ≡ PREVER
TRATAR-DE-1 (ús impersonal) Descripció-empresa
TRATAR-DE-2 (ús impersonal) Descripció-gestió
UNIR Fusió / Integració
VENDER-1 Resum-gestió
VENDER-2 Compra-venda (Si és una nominalització, en singular)

RELACIÓ INVERSA PER TIPUS DE VERBS

ACTIVITAT-EMPRESA

comercializar
exportar
fabricar
producir-2

ADMINISTRATIVA

acogerse-a
acordar
adoptar-2
alcanzar-2

aprobar
celebrar
confirmar-1
conseguir-2
decidir
esperar-1
estar-pendiente-de
firmar
llegar-a-2
obtener-2
pedir-2
proponer
ratificar
recibir-2
rematar
solicitar
suscribir-1

AMPLIACIÓ-DE-CAPITAL

ampliar-1
aumentar-2
elevar-1

COMANDA

pedir-1

COMPRA-VENDA

adquirir
ampliar-3
comprar-1
incorporar
ofrecer
poner-a-la-venta
vender-2

CREACIÓ

constituir-1

crear

DESCRIPCIÓ-EMPRESA

acoger

adoptar-1

considerar-1

constituir-3

contar-con

dar-empleo

emplear-a

formar

inscribir

ocupar

ocupar-a

operar-en

ostentar

poseer

ser-1

situarse-1

tener-2

tratar-de-1

DESCRIPCIÓ-GESTIÓ

afectar

considerar-2

constituir-2

pertenecer

responder-a

ser-2

ser-fruto-de

ser-resultado-de

sumar-1

tratar-de-2

DISCURSIVA

anticipar
anunciar
asegurar-1
confirmar-2
dar-a-conocer
esperar-2
estudiar
hacer-público
informar-1
revelar-1
señalar

DESTRUCTIVA

informar-2
parecer
rumorear

FUSIÓN

fusionar
incorporarse
unir

GESTIÓN-MÍNIMA

corresponder-a
crecer-2
desembolsar
emitir
equivaler
instalar
invertir-1
participar
pasar de ... a ...
quedar-fijado

recibir-1
reestructurar
renegociar
representar-2
ser-superior-a
superar
suponer-2
suscribir-2

INTEGRACIÓ

incorporarse
integrarse
unir

OPERADOR-RESUM-GESTIÓ

aumento
crecimiento
incremento

OPERADOR-VERBAL-BUIT

efectuar
hacer
llevar-a-cabo
realizar

OPERADOR-VERBAL-QUALIFICADOR

comprometerse
entrar-en
prever
proponerse
querer
tener-intención-de
tener-previsto

RESULTATIVA

ampliar-2
asegurar-2
consolidar
continuar-en
convertir-en
expandirse
introducirse
mejorar
modernizar
poner-de-relieve
representar-1
revelar-2
suponer-1

RESUM-GESTIÓ

acumular
alcanzar-1
ascender
aumentar-1
comprar-2
conseguir-1
crecer-1
descender
desinvertir
distribuir
duplicar
elevar-2
elevarse
estimar
facturar
incrementar
incrementarse
invertir-2
llegar-a-1

obtener-1
perder
producir-1
registrar
repartir
situarse-2
sumar-2
tener-1
vender-1

TANCAMENT

cerrar
declararse-en-bancarrota

Bibliografia

- [Allen-81] Allen, J.F. "An Interval-Based Representation of Temporal Knowledge". Actes 7è IJCAI, pàg. 221–226.
- [Allen-83] Allen, J.F. "Maintaining Knowledge about Temporal Intervals". *Communications of the ACM*, vol. 26, n. 11, pàg. 832–843.
- [Allen-84] Allen, J.F. "Towards a General Theory of Action and Time". *Artificial Intelligence*, vol. 33, n. 2, pàg. 123–154.
- [Allen-87] Allen, J.F. *Natural Language Understanding*. The Benjamin/Cummings Publishing Company.
- [AllHay-85] Allen, J. i Hayes, P.J. "A Common-Sense Theory of Time". Actes 9è IJCAI, pàg. 528–531.
- [AllHay-87] Allen, J. i Hayes, P.J. "Moments and Points in an Interval-Based Temporal Logic". TR 180 Depts. Computer Science and Philosophy. University of Rochester.
- [Appelo-86] Appelo, L. "A Compositional Approach to the Translation of Temporal Expressions in the Rosetta System". Actes COLING-86, pàg. 313–318.
- [ArToJa-85] Arnold, D.J., des Tombe, L. i Jaspert, L. "ELS-3 Eurotra Linguistic Specification. Version 3". Commission of European Communities, Luxembourg.
- [BoBoBras-88] Borillo, A., Borillo, M. i Bras, M. "A Temporal Reasoning Cognitive Approach". Actes de Workshop on Lexical Semantics and Artificial Intelligence (P. Saint-Dizier, ed.), Toulouse, pàg. 129–151.
- [BoBoBras-89] Borillo, A., Borillo, M. i Bras, M. "Sémantique du temps dans la phrase et dans le discours: de la théorie aux contraintes liées à l'implémentation. Présentat a la reunió sobre Sémantique logique du Temps, de l'Espace et du Mouvement dans le Langage Naturel, Toulouse, maig 1989.
- [Bruce-72] Bruce, B.C. "A Model for Temporal References and its Application in a Question Answering Program". *Artificial Intelligence*, vol. 3, n. 1, pàg. 1–25.
- [Bruce-73] Bruce, B.C. "The Processing of Time Phrases in Chronos". Technical Report CBM-TM-29. Laboratory for Computer Science Research, Rutgers University.
- [Castell-87] Castell, N. "Integrating Temporal Reasoning in a Frame-Based Formalism". Actes de International Conference on Artificial Intelligence. II World Basque Congress, pàg. 144–157. També a *Perspectives in Artificial Intelligence*, vol. 2. pàg. 86–96. Ellis Horwood Limited (1989).

- [Castell-88a] Castell, N. "Comprensió automàtica de notícies econòmiques". Actes IBERAMIA-88, pàg. 299-316.
- [Castell-88b] Castell, N. "Temporal Information Processing in the COTEM system". Actes de Workshop on Lexical Semantics and Artificial Intelligence (P. Saint-Dizier, ed.), Toulouse, pàg. 92-113.
- [CaMaVe-89a] Castell, N., Martí, M.A. i Verdejo, M.F. "Expression of time in news. Interpretation and representation". Pendent de publicació a un volum de la col·lecció *Langue, raisonnement et calcul*. Editors: M. Borillo i F. Nef.
- [CaMaVe-89b] Castell, N., Martí, M.A. i Verdejo, M.F. "Estudi lingüístic sobre l'expressió del temps a la narració periodística: notícies sobre economia". Report del Dep. LSI - Univ. Politècnica de Catalunya (pendent de publicació).
- [Colmerauer-79] Colmerauer, A. "Un sous-ensemble intéressant du français". *RAIRO Informatique*, vol. 13, n. 4, pàg. 1-28.
- [DeJong-79] DeJong, G.F. "Skimming Stories in Real Time: An Experiment in Integrated Understanding". Research Report #158. Dept. of Computer Science, Yale University.
- [DiRoVe-89] Díaz, A., Rodríguez, H. i Verdejo, M.F. "Un mecanismo para el tratamiento de la elipsis en sistemas dialogados". Congrès anual AEPIA-89.
- [Eynde-87] van Eynde, F. "Tense and Aspect". Versió preliminar d'un capítol del Reference Manual 4.0. Informe intern d'EUROTRA.
- [Eynde-88] van Eynde, F. "The Analysis of Tense and Aspect in EUROTRA". Actes COLING-88, pàg. 699-704.
- [FikKeh-85] Fikes, R. i Kehler, T. "The role of Frame-Based Representation in Reasoning". *Communications of the ACM*, vol. 28, n. 9, pàg. 904-920.
- [Hornstein-77] Hornstein, N. "Towards a Theory of Tense". *Linguistic Inquiry*, vol. 8, n. 3, pàg. 521-557.
- [IsaIshi-86] Isahara, H. i Ishizaki, S. "Context Analysis for Japanese Text". Actes COLING-86, pàg. 244-246.
- [KahGor-77] Kahn, K. i Gorry, G.A. "Mechanizing Temporal Knowledge". *Artificial Intelligence*, vol. 9, n. 2, pàg. 87-108.
- [Kamp-81] Kamp, H. "Événements, représentations discursives et référence temporelle". *Langages*, no. 64, pàg. 39-64.
- [KamRoh-83] Kamp, H. i Rohrer, C. "Tense in texts". *A Meaning, use and interpretation of language*, de Gruyter (Berlin), pàg. 250-269.

- [Kandrashina-83] Kandrashina, E.Yu. "Representation of Temporal Knowledge". Actes 8è IJCAI, pàg. 346–348.
- [Ladkin-86] Ladkin, P.B. "Time Representation: A Taxonomy of Interval Relations". Actes AAAI-86, pàg. 360–366.
- [LeDoFor-86] Leban, B., McDonald, D.D. i Forster, D.R. "A representation for Collections of Temporal Intervals". Actes AAAI-86, pàg. 367–371.
- [McDermott-82] McDermott, D. "A Temporal Logic for Reasoning About Processes and Plans. *Cognitive Science* vol. 6, pàg. 101–155.
- [Minsky-75] Minsky, M. "A Framework for Representing Knowledge". A *The Psychology of Computer Vision*. McGraw-Hill.
- [Obermeier-85a] Obermeier, K.K. "GROK – A Knowledge-Based Text Processing System". Actes Artificial Intelligence Applications Second Conference (IEEE), pàg. 384–389.
- [Obermeier-85b] Obermeier, K.K. "Temporal Inferences in Medical Texts". Actes 23è Annual Meeting of ACL, pàg. 9–17.
- [Reichenbach-47] Reichenbach, H. *The Elements of Symbolic Logic*. Macmillan, New York.
- [Rodríguez-89] Rodríguez, H. "GUAI: Un generador automático de interfaces en lengua natural". Tesi doctoral. Dep. LSI – Univ. Politècnica de Catalunya.
- [Sager-81] Sager, N. *Natural Language Information Processing: A Computer Grammar of English and its Applications*. Addison-Wesley.
- [SaFriLy-87] Sager, N., Friedman, C. i Lyman, N. *Medical Language Processing*. Addison-Wesley.
- [Schank-72] Schank, R.C. "Conceptual Dependency: a Theory of Natural Language Understanding". *Cognitive Psychology* vol. 3, pàg. 552–631.
- [SchaAbe-77] Schank, R.C. i Abelson, R.P. *Scripts, Plans, Goals and Understanding. An Inquiry into Human Knowledge Structures*. Lawrence Erlbaum Associates.
- [Vilain-82] Vilain, M.B. "A System for Reasoning about Time". Actes AAAI-82, pàg. 197–201.
- [VilKau-86] Vilain, M.B. i Kautz, H. "Constraint Propagation Algorithms for Temporal Reasoning". Actes AAAI-86, pàg. 377–382.
- [Yip-85] Yip, K.M. "Tense, Aspect and the Cognitive Representation of Time". Actes 10è IJCAI, pàg. 806–814.