A Learner with a Sense for Quality

Udo Hahn and **Klemens Schnattinger**

Text Knowledge Engineering Lab, Freiburg University
Werthmannplatz 1, D-79085 Freiburg, Germany
`http://www.coling.uni-freiburg.de`

Abstract

A text understanding system with learning capabilities is presented. New concepts are acquired by incorporating two kinds of evidence – knowledge about linguistic constructions in which unknown lexical items occur and knowledge about structural patterns in ontologies such that new concept descriptions can be compared with prior knowledge. On the basis of the quality of evidence gathered this way concept hypotheses are generated, ranked according to plausibility, and the most credible ones are selected for assimilation into the domain knowledge base.

Introduction

We propose a text understanding approach in which continuous enhancements of domain knowledge bases are performed given a core ontology (such as WordNet (Fellbaum, 1998)). New concepts are acquired taking two sources of evidence into account: the prior knowledge of the domain the texts are about, and linguistic constructions in which unknown lexical items occur. Domain knowledge serves as a comparison scale for judging the plausibility of newly derived concept descriptions in the light of prior knowledge. Linguistic knowledge helps to assess the strength of the interpretative force that can be attributed to the grammatical construction in which a new lexical item occurs. Our model makes explicit the kind of quality-based reasoning that lies behind such a process.

We advocate a *knowledge-intensive* model of concept learning from sparse data that is tightly integrated with the non-learning mode of text understanding. Both learning and understanding build on a given core ontology in the format of terminological assertions, and hence make abundant use of terminological reasoning facilities. The "plain" text understanding mode can be considered as the instantiation and continuous filling of roles with respect to *single concepts* already available in the knowledge base. Under learning conditions, a *set of alternative concept hypotheses* are managed for each unknown item, with each hypothesis denoting a newly created conceptual interpretation tentatively associated with the unknown item.

A Model of Quality-Based Learning

Fig. 1 depicts how linguistic and conceptual evidence are generated and combined for continuously discriminating

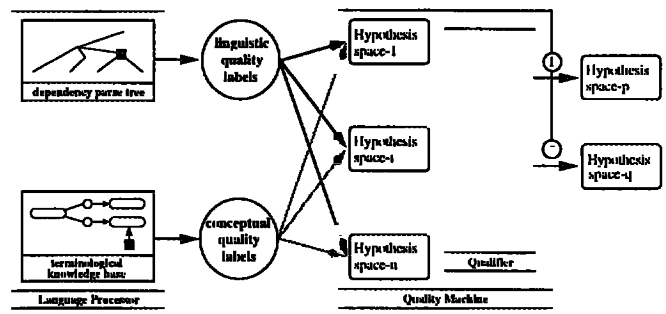

Figure 1: Architecture for Quality-Based Learning

and refining the set of concept hypotheses (the unknown item yet to be learned is characterized by the black square). The language processor yields structural dependency information from the grammatical constructions in which an unknown lexical item occurs in terms of the corresponding *parse tree*. The conceptual interpretation of parse trees involving unknown lexical items is used to derive *concept hypotheses*, which are further enriched by conceptual annotations reflecting structural patterns of consistency, mutual justification, analogy, etc. in the continuously updated *terminological knowledge base*. These kinds of initial evidence, in particular their predictive "goodness" for the learning task, are represented by corresponding sets of *linguistic* and *conceptual quality labels*. Multiple concept hypotheses for each unknown lexical item are organized in terms of a corresponding *hypothesis space*, each subspace holding different or further specialized concept hypotheses.

The *quality machine* estimates the overall credibility of single concept hypotheses by taking the available set of quality labels for each hypothesis into account. The final computation of a preference order for the entire set of competing hypotheses takes place in the *qualifier*, a terminological classifier extended by an evaluation metric for quality-based selection criteria. The output of the quality machine is a ranked list of concept hypotheses. The ranking yields, in decreasing order of significance, either the most plausible concept classes which classify the considered instance or more general concept classes subsuming the considered concept class.

Linguistic Quality Labels

Linguistic quality labels reflect structural properties of phrasal patterns or discourse contexts in which unknown lexical items occur – we assume here that the type of grammatical construction exercises a particular interpretative force on the unknown item and, at the same time, yields a particular level of credibility for the hypotheses being derived thereof. As an example of a high-quality label, consider the case of APPOSITION. This label is generated for constructions such as ".. *the printer @A@* ..", with "@..@" denoting the unknown item. The apposition almost unequivocally determines "@A@" (considered as a potential noun)[1] to denote an instance of the concept class PRINTER. This assumption is justified independent of further conceptual conditions, simply due to the nature of the linguistic construction being used. Still of good quality, but less constraining, are occurrences of the unknown item in a CASE-FRAME construction as illustrated by ".. *@B@ has a size of* ..". Here, case frame specifications of the verb "*has*" that relate to its AGENT role carry over to "*@B@*". Given its final semantic interpretation, "@B@" may be anything that has a size. Hence, considering an utterance like "*The Itoh-Ci-8 has a size of* ..", we may hypothesize that the concept ITOH-CI-8 can tentatively be considered a PRODUCT.

Let us now turn to a discussion of the phrase "*The switch of the Itoh-Ci-8* ..". We use a concept description language (for a survey, cf. Woods & Schmolze (1992)) for representing the content of texts and the emerging concept hypotheses. Considering this phrase, a straightforward translation into corresponding concept descriptions yields:

(P1) *switch*-01 : SWITCH
(P2) *Itoh-Ci*-8 HAS-SWITCH *switch*-01
(P3) HAS-SWITCH $\doteq$
(OUTPUTDEV $\sqcup$ INPUTDEV $\sqcup$ STORAGEDEV $\sqcup$ COMPUTER) $|$HAS-PART$|$ SWITCH

Assertion P1 indicates that the instance *switch*-01 belongs to the concept class SWITCH, P2 relates *Itoh-Ci*-8 and *switch*-01 via the binary relation HAS-SWITCH. The relation HAS-SWITCH is defined, finally, as the set of all HAS-PART relations which have their domain restricted to the disjunction of the concepts OUTPUTDEV, INPUTDEV, STORAGEDEVices or COMPUTER and their range restricted to SWITCHes.

Depending on the type of the syntactic construction in which the unknown lexical item occurs, different hypothesis generation rules may fire. In our example, "*The switch of the Itoh-Ci-8* ..", a genitive noun phrase places only few constraints on the item to be acquired. In the following, let *target* be the unknown item ("*Itoh-Ci-8*") and *base* be the known item ("*switch*"), whose conceptual relation to the target is constrained by the syntactic relation in which their lexical counterparts co-occur. The main constraint for genitives says that the target concept fills (exactly) one of the *n* roles of the base concept. Since the correct role cannot be yet decided upon, *n* alternative hypotheses have to be posited (unless additional constraints apply), and the target concept has to be assigned as a filler of the *i*-th role of base in the corresponding *i*-th hypothesis space. As a consequence, the classifier is able to derive a suitable concept hypothesis by specializing the target concept (initially TOP, by default) according to the value restriction of the base concept's *i*-th role. Additionally, this rule assigns a syntactic quality label to each *i-th* hypothesis indicating the type of syntactic construction in which target and base co-occur.

Returning to our example, the target concept ITOH-CI-8 is already predicted as a PRODUCT according to the interpretation of the phrase "*The Itoh-Ci-8 has a size of* ..". The conceptual representation of PRODUCT is given by:

PRODUCT $\doteq$
$\forall$HAS-PART.PHYSICALOBJECT $\sqcap$ $\forall$HAS-SIZE.SIZE $\sqcap$
$\forall$HAS-PRICE.PRICE $\sqcap$ $\forall$HAS-WEIGHT.WEIGHT

This expression reads as "all fillers of HAS-PART, HAS-SIZE, HAS-PRICE, HAS-WEIGHT roles must be concepts subsumed by PHYSICALOBJECT, SIZE, PRICE, WEIGHT, respectively". Accordingly, four roles remain to be considered for relating the target ITOH-CI-8 – as a tentative PRODUCT – to the base concept SWITCH. Three of them, HAS-SIZE, HAS-PRICE, and HAS-WEIGHT, are ruled out due to the violation of a simple integrity constraint (SWITCH does not denote a unit of measure). Therefore, only the role HAS-PART must be considered. Due to the definition of HAS-SWITCH (cf. P3), the instantiation of HAS-PART is specialized to HAS-SWITCH by the classifier, since the range of the HAS-PART relation is already restricted to SWITCH. Hence, four distinct hypotheses are immediately created due to the domain restrictions of the role HAS-SWITCH, *viz.* OUTPUTDEV, INPUTDEV, STORAGEDEV and COMPUTER, and are managed in four hypothesis spaces h_1, h_2, h_3 and h_4, respectively. We roughly sketch their contents in the following concept descriptions (note that for *Itoh-Ci*-8 we also include parts of the implicit *is-a* hierarchy):

$(Itoh\text{-}Ci\text{-}8 : \text{OUTPUTDEV})_{h_1}, (Itoh\text{-}Ci\text{-}8 : \text{DEVICE})_{h_1}, ..,$
$(Itoh\text{-}Ci\text{-}8\ \text{HAS-SWITCH}\ Switch.0\text{-}00025)_{h_1}$

$(Itoh\text{-}Ci\text{-}8 : \text{INPUTDEV})_{h_2}, (Itoh\text{-}Ci\text{-}8 : \text{DEVICE})_{h_2}, ..,$
$(Itoh\text{-}Ci\text{-}8\ \text{HAS-SWITCH}\ Switch.0\text{-}00025)_{h_2}$

$(Itoh\text{-}Ci\text{-}8 : \text{STORAGEDEV})_{h_3}, (Itoh\text{-}Ci\text{-}8 : \text{DEVICE})_{h_3}, ..,$
$(Itoh\text{-}Ci\text{-}8\ \text{HAS-SWITCH}\ Switch.0\text{-}00025)_{h_3}$

$(Itoh\text{-}Ci\text{-}8 : \text{COMPUTER})_{h_4}, (Itoh\text{-}Ci\text{-}8 : \text{HARDWARE})_{h_4}, ..,$
$(Itoh\text{-}Ci\text{-}8\ \text{HAS-SWITCH}\ Switch.0\text{-}00025)_{h_4}$

Conceptual Quality Labels

Conceptual quality labels result from comparing the representation structures of a concept hypothesis with already existing representation structures in the underlying domain knowledge base from the viewpoint of structural similarity, incompatibility, etc. The closer the match, the more credit is lent to a hypothesis. For instance, a very positive conceptual quality label such as M-DEDUCED is assigned to multiple derivations of the same concept hypothesis in different hypothesis (sub)spaces. Positive labels are also assigned to terminological expressions which share structural similarities, though they are not identical. For instance, the label C-SUPPORTED is assigned to any hypothesized relation $R1$ between two instances when another relation, $R2$, already exists in the KB involving the same two instances,

[1] Such a part-of-speech hypothesis can directly be derived from the inventory of valence and word order specifications underlying the dependency grammar model we use (Hahn et al., 1994).

but where the role fillers occur in "inverted" order (note that $R1$ and $R2$ need not necessarily be semantically inverse relations such as with *"buy"* and *"sell"*). This rule of cross support captures the inherent symmetry between concepts related via quasi-inverse conceptual relations.

Considering our example, for ITOH-CI-8 the concept hypotheses OUTPUTDEV, INPUTDEV and STORAGEDEV were derived independently of each other in different hypothesis spaces. Hence, DEVICE as their common superconcept has been multiply derived by the classifier in each of these spaces, too. Accordingly, this hypothesis is assigned a high degree of confidence by the classifier which derives the conceptual quality label M-DEDUCED:

$$(Itoh\text{-}Ci\text{-}8 : \text{DEVICE})_{h_1} \sqcap (Itoh\text{-}Ci\text{-}8 : \text{DEVICE})_{h_2} \Longrightarrow$$
$$(Itoh\text{-}Ci\text{-}8 : \text{DEVICE})_{h_1} : \text{M-DEDUCED} \ \ldots \ \ldots$$

Quality-Based Classification

Whenever new evidence for or against a concept hypothesis is brought forth in a single learning step all concept hypotheses are reevaluated. First, weak or even untenable hypotheses are eliminated from further consideration. The corresponding quality-based selection among hypothesis spaces is grounded on *threshold levels* (later on referred to as **TH**). Their definition takes mostly linguistic evidence into account. At the first threshold level, all hypothesis spaces with the maximum of APPOSITION labels are selected. If more than one hypothesis is left to be considered, only concept hypotheses with the maximum number of CASEFRAME assignments are approved at the second threshold level. Those hypothesis spaces that have fulfilled these threshold criteria will then be classified relative to two different *credibility levels* (later on referred to as **CB**). The first level of credibility contains all hypothesis spaces which have the maximum of M-DEDUCED labels, while at the second level (again, with more than one hypothesis left to be considered) those are chosen which are assigned the maximum of C-SUPPORTED labels. A more technical terminological specification of the entire qualification calculus is given by Schnattinger & Hahn (1996).

For an illustration, consider the first phrase: *"The Itoh-Ci-8 has a size of.."*. The assignment of the syntactic quality label CASEFRAME to this phrase is triggered only in those hypothesis spaces where the unknown item is considered a PHYSICALOBJECT (cf. Table 2, learning step 1). The remaining hypotheses (cf. Table 2, learning step 2) cannot be annotated by CASEFRAME, since the concepts they represent have no property such as SIZE. As a consequence, their hypothesis spaces are ruled out by the criterion set up at the second threshold level, and the still valid concept hypothesis PHYSICALOBJECT is further refined as PRODUCT. As far as the sample phrase *"The switch of the Itoh-Ci-8 .."* is concerned, four more specific hypothesis spaces are generated from the PRODUCT hypothesis, three of which stipulate a DEVICE hypothesis. Since the conceptual quality label M-DEDUCED has been derived by the classifier, this result yields a preliminary ranking with these three DEVICE hypotheses preferred over the one associated with COMPUTER (cf. Table 2, learning step 3).

Phrase	Semantic Interpretation
1. The *Itoh-Ci-8*	(possess.1,agent,Itoh-Ci-8)
2. *has* a *size* of ..	(possess.1,patient,Size.1)
	↦ (Itoh-Ci-8,has-size,Size.1)
3. The *switch* of the *Itoh-Ci-8*	(Itoh-Ci-8,has-switch,Switch.1)
4. The *housing* of the *Itoh-Ci-8*	(Itoh-Ci-8,has-housing,Housing.1)
5. *Itoh-Ci-8* with a *memory*	(Itoh-Ci-8,has-memory,Memory.1)
6. *Itoh-Ci-8's LED lines*	(Itoh-Ci-8,has-part,LED-Line.1)
7. *Itoh-Ci-8* with a *resolution*	(Itoh-Ci-8,has-rate,Resolution.1)

Table 1: Phrases and Interpretations Related to *"Itoh-Ci-8"*

Evaluation

In this section, we present data from an empirical evaluation of the text learner. We considered a total of 101 texts taken from a corpus of information technology magazines. For each of them 5 to 15 learning steps were considered. A *learning step* consists of the inferences being made at the level of hypothesis spaces after new textual input has been supplied in which the item to be learned occurs. In order to clarify the input data for the learner, consider Table 1. It consists of seven single phrases in which the unknown item *"Itoh-Ci-8"* occurs, together with the respective semantic interpretations. The knowledge base on which we performed our experiments currently comprises 325 concept definitions and 447 conceptual relations.

In a series of experiments, we investigated the learning accuracy of the system, i.e., the degree to which the system correctly predicts the concept class which subsumes or classifies the target concept to be learned. Taxonomic hierarchies emerge naturally in terminological knowledge representation frameworks. So a prediction can be more or less precise, i.e., it may approximate the goal concept at different levels of specificity. This is captured by our measure of *learning accuracy* which takes into account the conceptual distance of a hypothesis to the goal concept of an instance. Learning accuracy (LA) is defined here as (n being the number of concept hypotheses for a single target):

$$LA := \sum_{i \in \{1 \ldots n\}} \frac{LA_i}{n} \quad \text{with}$$

$$LA_i := \begin{cases} \frac{CP_i}{SP_i} & \text{if } FP_i = 0 \\ \frac{CP_i}{FP_i + DP_i} & \text{else} \end{cases}$$

SP_i specifies the length of the *shortest path* (in terms of the number of nodes being traversed) from the TOP node of the concept hierarchy to the maximally specific concept subsuming the instance to be learned in hypothesis i; CP_i specifies the length of the path from the TOP node to that concept node in hypothesis i which is *common* to both the shortest path (as defined above) and the actual path to the predicted concept (whether correct or not); FP_i specifies the length of the path from the TOP node to the predicted (in this case *false*) concept and DP_i denotes the node *distance* between the predicted (false) node and the most specific common concept (on the path from the TOP node to

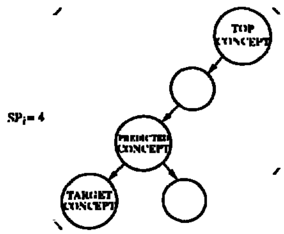

Figure 2: LA for an Underspecified Concept Hypothesis

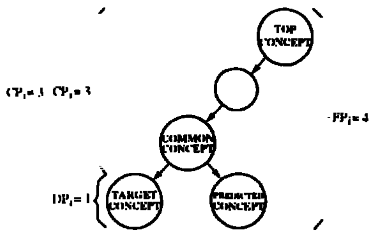

Figure 3: LA for a Slightly Incorrect Concept Hypothesis

the predicted false node) still correctly subsuming the target in hypothesis i. Figures 2 and 3 depict sample configurations for concrete LA values involving these parameters. Fig. 2 illustrates a correct, yet too general prediction with $LA_i = .75$, while Fig. 3 contains an incorrect concept hypothesis with $LA_i = .6$. Though the measure is sensitive to the depth of the concept graphs in a knowledge base, it produced adequate results in the information technology domain we considered. As the graphs in knowledge bases for "natural" domains typically have an almost canonical depth that ranges between seven to ten nodes from the most general to the most specific concept (cf., e.g., the *WordNet* lexical database (Fellbaum, 1998)), our experience seems to generalize to other domains as well.

Given this measure, Table 2 illustrates how the various concept hypotheses for Itoh-Ci-8 develop in learning accuracy from one step to the other, relative to the data from Table 1. The numbers in brackets in the column **Concept Hypotheses** indicate for each hypothesized concept the number of subsumed concepts in the underlying knowledge base; **LA CB** gives the accuracy rate for the full qualification calculus including threshold and credibility criteria, **LA TH** for threshold criteria only, while **LA –** depicts the accuracy values produced by the terminological reasoning component without incorporating these quality criteria. As can be seen from Table 2, the full qualification calculus produces either the same or even more accurate results, the same or fewer hypothesis spaces (indicated by the number of rows), and derives the correct prediction more rapidly (in step 6) than the less knowledgeable variants (in step 7).

The data also illustrate the continuous specialization of concept hypotheses achieved by the terminological classifier, e.g., from PhysicalObject[2] in step 1 via Product in step 2 to OutputDevice and Printer in steps 3 and 4, respectively. The overall learning accuracy may even temporarily decrease in the course of hypothesizing (e.g., from step 3 to 4 or step 5 to 6 for **LA –** and **LA TH**), but the learning accuracy value for the full qualification calculus (**LA CB**) always increases. Fig. 4 depicts the learning accuracy curve for the entire data set (101 texts). The evaluation starts from LA values in the interval between 48% to 54% for **LA –/LA TH** and **LA CB**, respectively, in the first learning step. In the final step, LA rises up to 79%, 83% and 87% for **LA –**, **LA TH** and **LA CB**, respectively.

The pure terminological reasoning machinery which does not incorporate the qualification calculus always achieves an inferior level of learning accuracy and generates more hypothesis spaces than the learner equipped with the qualification calculus. Furthermore, the inclusion of conceptual criteria (**CB**) supplementing the linguistic criteria (**TH**) helps a lot to focus on the relevant hypothesis spaces and to further discriminate the valid hypotheses (in the range of 4% precision). Note that a significant plateau of accuracy is usually reached after the third step (*viz.* 67%, 73%, and 76% for **LA –**, **LA TH**, and **LA CB**, respectively, in Fig. 4). This indicates that our approach finds the most relevant distinctions in a very early phase of the learning process, i.e., it requires only a *few* examples.

Concept Hypotheses	**LA –**	**LA TH**	**LA CB**
PhysicalObject(176)	0.30	0.30	0.30
MentalObject(0)	0.16	0.16	0.16
InformationObject(5)	0.16	0.16	0.16
MassObject(0)	0.16	0.16	0.16
Norm(3)	0.16	0.16	0.16
Technology(1)	0.16	0.16	0.16
Mode(5)	0.16	0.16	0.16
Feature(0)	0.16	0.16	0.16
Learning Step 1	ϕ:0.18	ϕ:0.18	ϕ:0.18
Product(136)	0.50	0.50	0.50
MentalObject(0)	0.16		
InformationObject(5)	0.16		
MassObject(0)	0.16		
Norm(3)	0.16		
Technology(1)	0.16		
Mode(5)	0.16		
Feature(0)	0.16		
Learning Step 2	ϕ:0.21	ϕ:0.50	ϕ:0.50
Computer(5)	0.50	0.50	
OutputDevice(9)	0.80	0.80	0.80
StorageDevice(5)	0.55	0.55	0.55
InputDevice(2)	0.55	0.55	0.55
Learning Step 3	ϕ:0.60	ϕ:0.60	ϕ:0.63
Notebook(0)	0.43	0.43	
Portable(0)	0.43	0.43	
PC(0)	0.43	0.43	
Workstation(0)	0.43	0.43	
Desktop(0)	0.43	0.43	
Printer(3)	0.90	0.90	0.90
VisualDevice(2)	0.66	0.66	0.66
Loudspeaker(0)	0.66	0.66	0.66
Plotter(0)	0.66	0.66	0.66
RW-Store(2)	0.50	0.50	0.50
RO-Store(1)	0.50	0.50	0.50
Mouse(0)	0.50	0.50	
Keyboard(0)	0.50	0.50	
Learning Step 4	ϕ:0.54	ϕ:0.54	ϕ:0.65
Notebook(0)	0.43	0.43	
Portable(0)	0.43	0.43	
PC(0)	0.43	0.43	
Workstation(0)	0.43	0.43	
Desktop(0)	0.43	0.43	
LaserPrinter(0)	1.00	1.00	1.00
InkjetPrinter(0)	0.75	0.75	0.75
NeedlePrinter(0)	0.75	0.75	0.75
Learning Step 5	ϕ:0.58	ϕ:0.58	ϕ:0.83
Notebook(0)	0.43	0.43	
Portable(0)	0.43	0.43	
PC(0)	0.43	0.43	
Workstation(0)	0.43	0.43	
Desktop(0)	0.43	0.43	
LaserPrinter(0)	1.00	1.00	1.00
Learning Step 6	ϕ:0.52	ϕ:0.52	ϕ:1.00
LaserPrinter(0)	1.00	1.00	1.00
Learning Step 7	ϕ:1.00	ϕ:1.0	ϕ:1.00

Table 2: Learning Steps for a Text Featuring *"Itoh-Ci-8"*

[2]The upper-level concepts of our domain ontology were taken from Nirenburg & Raskin (1987).

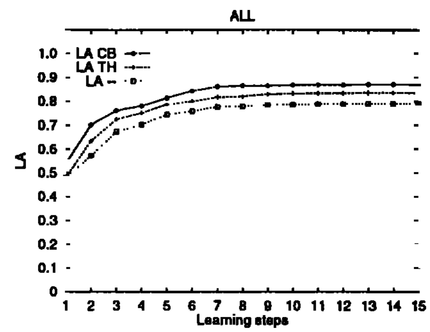

Figure 4: Learning Accuracy (LA) for the Entire Data Set

Related Work

Our approach bears a close relationship to the work of Mooney (1987), Gomez & Segami (1989), Rau et al. (1989), Hastings (1996), and Moorman & Ram (1996), who all aim at the automated learning of word meanings from context using a knowledge-intensive approach. Our work differs from theirs in that the need to cope with *several competing* concept hypotheses and to aim at a *reason-based selection* is not an issue in those studies.

The work closest to ours has been carried out by Rau et al. (1989) and Hastings (1996). Concept hypotheses are also generated from linguistic and conceptual data. Unlike our approach, the selection of hypotheses depends only on an ongoing discrimination process based on the availability of these data but does not incorporate an inferencing scheme for reasoned hypothesis selection. The difference in learning performance – in the light of our evaluation study – amounts to 8%, considering the difference between **LA -** (plain terminological reasoning) and **LA CB** values (terminological metareasoning based on the qualification calculus). Hence, our claim that we produce competitive results.

Note that the requirement to provide learning facilities for large-scale text understanders also distinguishes our approach from the currently active field of information extraction (IE) (Appelt et al., 1993). The IE task is defined in terms of a *pre-fixed* set of templates which have to be instantiated (i.e., filled with factual knowledge items) in the course of text analysis. Unlike the procedure we propose, no *new* templates have to be created.

Conclusion

In this paper, we have introduced a methodology for generating new knowledge items from texts and integrating them into an existing domain knowledge base. This is based on the incremental assignment and evaluation of the quality of linguistic and conceptual evidence for emerging concept hypotheses. The concept acquisition mechanism we propose is fully integrated in the text understanding mode. No specialized learning algorithm is needed, since learning is a (meta)reasoning task carried out by the classifier of a terminological reasoning system. However, heuristic guidance for selecting between plausible hypotheses comes from the different quality criteria. Our experimental data indicate that given these heuristics we achieve a high degree of pruning of the search space for hypotheses in very early phases of the learning cycle.

Our experiments are still restricted to the case of a single unknown concept in the entire text. Generalizing to n unknown concepts can be considered from two perspectives. When hypotheses of another target item are generated and incrementally assessed relative to an already given base item, no effect occurs. When, however, two targets (i.e., two unknown items) have to be related, then the number of hypotheses that have to be taken into account is equal to the product of the number of hypothesis spaces currently associated with each of them. In the future, we intend to study such scenarios. Fortunately, our evaluation results also indicate that the number of hypothesis spaces decreases rapidly as does the learning rate, i.e., the number of concepts included in the remaining concept hypotheses. So, the learning system should remain within feasible bounds, even under these less favorable conditions.

Acknowledgements. We would like to thank our colleagues in the CLIF group for fruitful discussions and instant support. K. Schnattinger is supported by a grant from DFG (Ha 2097/3-1).

References

Appelt, D., J. Hobbs, J. Bear, D. Israel & M. Tyson (1993). FASTUS: A finite-state processor for information extraction from real-world text. In *IJCAI'93 – Proc. 13th Intl. Conf. on Artificial Intelligence*, pp. 1172–1178.

Fellbaum, C. (Ed.) (1998). *WordNet: An Electronic Lexical Database.* Cambridge, MA: MIT Press.

Gomez, F. & C. Segami (1989). The recognition and classification of concepts in understanding scientific texts. *J. of Exp. and Theoret. Artif. Intelligence*, 1:51–77.

Hahn, U., S. Schacht & N. Bröker (1994). Concurrent, object-oriented natural language parsing: the PARSETALK model. *International Journal of Human-Computer Studies*, 41(1/2):179–222.

Hastings, P. (1996). Implications of an automatic lexical acquisition system. In S. Wermter, E. Riloff & G. Scheler (Eds.), *Connectionist, Statistical and Symbolic Approaches to Learning in Natural Language Processing*, pp. 261–274. Berlin: Springer.

Mooney, R. (1987). Integrated learning of words and their underlying concepts. In *CogSci'87 - Proc. 9th Annual Conf. of the Cognitive Science Society*, pp. 974–978.

Moorman, K. & A. Ram (1996). The role of ontology in creative understanding. In *CogSci'96 – Proc. 18th Annual Conf. of the Cognitive Science Soc.*, pp. 98–103.

Nirenburg, S. & V. Raskin (1987). The subworld concept lexicon and the lexicon management system. *Computational Linguistics*, 13(3-4):276–289.

Rau, L., P. Jacobs & U. Zernik (1989). Information extraction and text summarization using linguistic knowledge acquisition. *Information Processing & Management*, 25(4):419–428.

Schnattinger, K. & U. Hahn (1996). A sketch of a qualification calculus. In *FLAIRS'96 – Proc. 9th Florida AI Research Symposium*, pp. 198–203.

Woods, W. & J. Schmolze (1992). The KL-ONE family. *Computers & Math. with Applications*, 23:133–177.