Abstract
The use of default logic to represent various linguistic constructions is explored in this paper. Default logic is then integrated into a theory of natural language semantics, namely Generalized Quantifiers. Finally, properties of interest to the AI community such as the characterization of truth persistence and some inferential patterns are emphasized.

1 Introduction
Default logic [Reiter 80] is now commonly used in artificial intelligence systems and turns out to be one of the major research fields in knowledge representation. One of the first motivations of default logic was to give a more accurate semantic representation to natural language statements of the form:

-Birds fly.

Most birds fly.

or arguably:

Typically, birds fly.

The idea was to define a representation that permits exceptions to be coherent together with a general statement or law. Very soon this linguistic motivation became somewhat neglected, yielding the way to theoretical investigations in non-monotonic reasoning and to the elaboration of automatic theorem provers. For a good introduction see [Besnard 89]. The only works we are aware of in natural language processing are the use of default logic to solve anaphoras [Dunin-Replita 84], to model dialogs [Joshi et al. 87] and to derive presuppositions from sentences [Mercer 88].

In this contribution, we come back to the very first motivation of default logic and explore its integration into natural language semantics, as introduced in [Saint-Dizier 87]. More precisely, we propose an integration of default logic into Generalized Quantifiers framework [Barwise and Cooper 81]. This leads us to reformulate default logic within the framework of set theory defined on universes. We first propose several types of linguistic expressions for which a representation by a default rule turns out to be relevant. Next, we present some basic properties about truth persistence of statements represented by default logic and, finally, we propose several inferential patterns which permit to derive new default rules on the one hand and new linguistic expressions on the other hand.

2 Default logic for natural language semantics
2.1 Introduction to default logic
A default rule has the following general form:

\[ P : R \]

\[ C \]

The meaning of this deduction rule is that given \( x \) such that \( P \) is true, if there are no contradiction to \( R \) being true in the current world, then infer \( C \). \( P \) is called the prerequisite and \( C \) is the conclusion. \( R \) is a condition which has to be coherent with the current world.

Such a rule can be used to represent general laws or statements such as:

Most ravens are black.

or:

All ravens are black.

as well as contingent facts such as:

Most lights are on.

The first sentence above is represented as:

\[ \text{raven}(X) : \text{black}(X) \]

\[ \text{black}(X) \]

(variables are represented by capital letters)

The difference between general laws and contingent statements is however fundamental but it cannot be captured straightforwardly within default logic. Indeed, in the first example, blackness is taken as a property inherent to ravens, not completely based on observation but also on induction. Default logic permits this property to remain inherent to ravens even if there are exceptions. The second sentence is strictly based on observation and is valid only on a certain time interval \( t \) at a given location \( l \). The distinction between contingent facts and general laws can however be context-dependent. Suppose the above sentence is stated in the context of a factory working 24 hours every day then the fact that lights are on can be interpreted as a law imposed, for instance, for security reasons.

There is another difference: the first sentence may be true even if there are no ravens in the current world whereas the second sentence requires a non-empty denotation for the part of noun phrase following most, in other terms, the presence in our example of at least two lights.

Default logic permits to assign a property \( (C) \) to a precise individual in a given world, modulo some coherence conditions \( (R) \). This can be contrasted with fuzzy and probabilistic logics where a set of individuals is characterized as a whole by a property. Thus, no property can be deduced for a given individual in this set with a full degree of certainty. Our choice of default logic is then motivated by the need of being able to make inferences about a precise individual and to provide precise responses. This does not exclude fuzzy and probabilistic logics, but it simply has different motivations and uses.
Using default logic also permits to skip over some recalcitrant representational problems of fuzzy logic, but, on the other hand, it cannot take into account slight differences between, for example:

Most birds fly versus Almost all birds fly.
A majority of birds fly versus A high majority of birds fly.

The question however is to know what is the real nature of the difference between those statements and if it is more than a connotative meaning.

At a formal level, it turns out that additional inferential patterns can be formulated when using default reasoning. Furthermore, due to a certain stability, a default rule used in a semantic representation of a sentence conveys to this representation useful properties in a knowledge base context such as: conservativity, extensibility and some forms of monotonicity. Because of their generality and stability, default rules can also play a prominent role to restore consistency in knowledge bases.

2.2 Classes of words and constructions represented by default logic

In a sentence, a wide diversity of types of words and linguistic constructions confer a certain degree of generality (either contingent or permanent) to the statement in which they are included. In this class fall words like context-dependent determiners, quantification introducing distributivity: where the context-dependent determiner is in the scope of a logic. By relative meaning, we mean, for example, sentences a relative meaning cannot in general be represented by default logic is to substitute all or no respectively for most, many ... and few and to check if it is possible to build a coherent world from the current world with less exceptions (i.e. a more uniform world as in [Delgrande 87]) so that the universally quantified statement is true. If there is such a world then the original determiner can be represented by a default rule.

For example, consider the sentence:

John met many people to-day.

Considered in a world where John is the manager of a company, people refers to the staff of the company and John met every-body in the company except n people (n being small). If we substitute all for many, we obtain the sentence:

John met all people to-day.

If if is a politician and people refers to the people who want to vote for him (suppose this number is large), then it is not possible to consistently build a world in which John met every person in a single day, in particular if those people are scattered throughout the whole country. In this latter case, many cannot be represented by a default rule.

2.2.2 Quantificational adverbs

Adverbs introducing a form of universality or contributing to it can originate a default logic based representation:

John sings rarely.
Mary is incidentally (very rarely) on time.
John meets Sue almost every day.
John often travels by bus rather than by subway in Paris.

For default logic to be used, the concept or the quality modified by the adverb has to be countable and quantifiable. The first example can be paraphrased as:

At most times t (or preferably, on most time intervals t), by default John does not sing.

The representation of sentence 2 requires the use of events and is far more complex. Roughly speaking, we can paraphrase this sentence by:

In most cases when Mary has something to do, she does it later than the scheduled time for this work.

This can be formulated in an equivalent way in terms of default rule. If e1 is the event associated with the scheduled time for the beginning of the activity and if e2 is the event of Mary beginning to do this activity, then if there is no contradiction with e2 > e1 then infer that e2 > e1.

Sentence 3 requires events to be associated with, for example, a date and a duration. It can be paraphrased by:

If it is not inconsistent that John meets Sue to-day (e.g. John is in the same location as Sue, etc...), then infer that there is a time interval during which John meets Sue.

Various factors like temporal factors can restrict the scope of a default:

From 8 to 10 am, Mary is very rarely in her office.
but they still have a universal meaning over that restricted domain.

2.2.3 Adjectives

Adjectives in the superlative, in particular when the statement is not absolutely incompatible with a small number of exceptions can be represented by a default rule. This permits to avoid a too strong and definitive formulation. Consider:

The Mon-BLANC is (one of) the highest mountains.

Representing this sentence by default logic permits:

- to assume that a mountain whose elevation is unknown is lower than the Mon-BLANC,
- to accept without introducing inconsistencies that there are mountains with an elevation explicitly known and greater than the Mon-BLANC elevation.

This approach is particularly relevant for superlatives: it turns out indeed that most superlatives are not completely universal. They are often true in a coherent subset of the current world.

In the same range of ideas, adjectives such as ideal, worse of perfect turn out to be, in most contexts, implicit superlatives ranging over several properties. In:

Jim is the ideal hiker.

ideal could mean that Jim has most of the qualities of a hiker in something close to the superlative. Thus, such a sentence can be represented by a conjunction of properties in the superlative, each of them being separately represented by a default rule.

The range of properties referred to by adjectives like ideal or worse can be restricted (or a specific property can be emphasized) as in:

Mary is the worst person to work with: she is (almost) never on time.

where the quality referred to is to be never on time. In this case, the representation becomes much simpler.

2.2.4 Agentless passive constructions

Agentless passive constructions of verbs like to admire, to hate, to neglect or to laugh at, in some contexts lend very well themselves to a representation with defaults. For example, a sentence like:

Mary is admired.

can be paraphrased by:

Most persons who know Mary admire her.

2.2.5 Propositional adverbs and adjectives

Adverbs and adjectives such as probable, likely, unlikely and certainly in some contexts modify the certainty of a statement, as in the sentence:

It is unlikely that computer science students know baroque music composers.

These adverbs and adjectives are basically interpreted with an intensional meaning. However, they introduce another interesting type of relation, out of the scope of the present contribution.

3 Integrating default logic into Generalized Quantifiers

In the examples we have given above, we have restricted our attention to the most straightforward uses of default logic. In particular, the expressions or concepts used are:

- completely defined (no pronominal references, ...),
- extensional,
- discrete (no continuous uses as in "a lot of snow").

None of these restrictions are, however, essential. Another more important restriction is that universes are supposed to be finite at each time t, in order to make our notations below computationally tractable.

3.1 Generalized Quantifiers

To deal with semantic representations, we adopt the spirit of the Generalized Quantifiers framework [Barwise and Cooper 81]. This approach is, in fact, of much interest because it is quite close to current research in knowledge bases. A generalized quantifier Q denotes a relation among sets of entities in a world W. It is noted as:

Q A B

where A and B are linguistic expressions or their set-theoretic equivalents. For example:

All ravens are black

is noted as:

All (ravens) are black.

All establishes a relation between the set of ravens and the set of entities which are black. Let us note the denotation of A in world W (|| A ||w) A' and that of B (|| B ||w), B'. Then, every generalized quantifier Q can be represented by a corresponding numerical relation RQ [Van Benthem 86], [Vesterstahl 84,85] with the following definition:

Q A B \iff RQ(|| A' \cap B' ||, || A' \cup B' ||).

In Generalized Quantifiers, determiners are studied in a principled way by looking at their semantic properties. This study appears to have enough logical foundations to motivate theoretical investigations. Generalized Quantifiers also turn out to be extended to represent determiners but extends to the semantics of other structures such as conditionals [Van Benthem 86].

We now turn to informally introduce default logic into the Generalized Quantifiers framework. A default rule is basically used to conclude a formula C for a given entity e, satisfying a prerequisite P, modulo R. However, it is also possible to characterize the set E of elements e satisfying P and for which C can be concluded modulo the coherence control on R. If we view a default rule as a ternary relation:

by-def P R C.

or, simply as a binary relation since, in our context, R and C are identical:

by-def P C.

then || P ||w and || C ||w can be defined and the above relation is a relation among sets of individuals over world W in a way similar to the determiners accounted for within the Generalized Quantifiers framework. In addition, a third set should be mentioned, which is the set of exceptions:

\epsilon = || P \land \neg C ||w.

which introduces another interesting type of relation, out of the scope of the present contribution.
3.2 Some examples

As an illustration, we now present possible representations for some of the examples given above in the previous section. Those representations have a knowledge-base orientation rather than a pure formal semantics one.

(a) Many workers have a car.
   is represented by:
   by-def(worker(X), car(Y) ∧ to-have(X, Y)).

(b) John knows few bird names.
   is represented by:
   by-def(bird(Y) ∧ name-of(X, Y),
   ¬ to-know(john, name-of(Y, X)), Y).

(c) John sings rarely.
   is represented by:
   by-def(time(T), ¬ to-sing(john, T)).
   T is a precise time or, preferably, a time interval.

(d) John meets Sue almost every day.
   is represented as follows:
   by-def(day(D), (3 H1, H2, H2 > H1,
   to-meet(john, sue, time(D, H1, H2)))).

(e) John often travels by bus rather than by subway.
   is represented by:
   by-def(bus(X) ∧ subway(Y) ∧ to-travel(john, Z) ∧ ((Z=X) ∨ (Z=Y)), Z=X).

This logical representation means that if John travels by Z which can be either a bus or a subway (X or Y), then, by default, John travels by bus (i.e. by X).

(f) The Mont-Blanc is one of the highest mountains.
   is represented as follows:
   by-def((mountain(X) ∧ ¬(X = mont-blanc) ∧ elevation(X, Y) ∧ elevation(mont-blanc, M)), Y < M).

(g) Mary is admired.
   is represented by:
   by-def(to-know(X, mary), to-admire(X, mary)).

We could also add that X has the ability to admire.

4 Stability of statements represented by default rules

By stability of a statement, we mean the characterization of conditions under which a statement remains true when the current world is updated. In our framework, stability means the characterization of the conditions under which the set of elements x that satisfy by-def A B remains unchanged, i.e. any deduction made from that default rule for any individual x before the updating remains true after the updating.

Representations with defaults appear to have slightly different properties than their more classical counterparts. Among those properties, we now present some of those which are of much interest to knowledge representation systems. The properties listed below are central to the field of Generalized Quantifiers.

4.1 Conservativity

For all A, B being linguistic expressions (or their set-theoretic equivalents):

by-def A B ↔ by-def A (A ∧ B) (noted CONS)

The equivalence is straightforward in virtue of the very nature of the prerequisite A.

4.2 Extension

Let us first introduce the notion of irrelevance. The idea is that propositions irrelevant to a default statement should have no consequence on any inference involving that statement. This idea was developed by the philosophical community and, more recently by [Delgrande 87]. In our framework, if we consider the statement: by-def A B, then, intuitively, C is irrelevant for this statement if knowing C does not alter the set of individuals for which that statement holds. We differentiate right and left irrelevance:

Let W, W’ and W” be worlds defined on universe U.

Let A, B and C be linguistic expressions or their set-theoretic equivalents. We consider the statement by-def A B . Then we have the following properties:

- left-irrelevance:
  C is left-irrelevant iff: W’ = W ∪ {C} then || A ||w = || A ||w’.

- right-irrelevance:
  Let W” = W ∪ {C’}, then C’ is right-irrelevant iff:
  \{ z ∈ W | consistentw(B(z..)) \} =
  \{ y ∈ W” | consistentw(B(y..)) \}.

consistentw(S) is a predicate which is true if in world W the statement S is consistent.

If W contains disjunctions of formulae, then it is necessary to consider all maximal extensions E of W to define right-irrelevance. The following condition has to be true:

For all E, maximal extension of W such that:

1. E” = E ∪ {C}.
2. E is consistent in W.
3. \{ x ∈ E | consistentw(B(z..)) \} =
   \{ y ∈ E” | consistentw(B(y..)) \}.

Then, we say that C is irrelevant to by-def A B if it is both left and right irrelevant. Then the property of extension follows:

For all A, B, for all W, W’ such that W’ is an extension of W where only irrelevant statements to by-def A B have been added, then:

by - defw A B ⇐⇒ by - defw’ A B. (noted EXT).

4.3 Monotonicity

For all A, B, W, W’ :

(a) Let W’ be an extension of W such that W ∩ W’ is a set of left-irrelevant statements then:
   by - defw’ A B ⇒ by - defw A B (downward left-monotonicity, noted ↓MON ).

(b) Let W” be an extension of W such that W ∩ W” is a set of right irrelevant statements then:
   by - defw” A B ⇒ by - defw A B (downward right-monotonicity, noted ↑MON ).

Not surprisingly (default logic is a non-monotonic logic), upward monotonicity does not hold in general.

5 Inferential patterns

Some inferential patterns within Generalized Quantifiers [Vesterstahl 84,85], [Van Benthem 86] also hold, with some restrictions, for default logic. Some additional patterns can be formulated, given the specificities of default logic. These patterns permit to derive new rules from previous ones and to generate new linguistic expressions. Here are some simple, basic inferential patterns:

5.1 Restricted transitivity

For all A, B, C linguistic expressions:

(by-def A (is B)) ∨ (all B) C ⇒ (by-def A C).

(is B), (all B), ... are meta-linguistic expressions corresponding to well-formed linguistic expressions explicitly con-
Note that in the patterns already stated, the determiner at the origin of the default rule has been weakened, i.e., on how much the number of exceptions to the default rule has increased. This number is characterized by the cardinal of the following set:

\[ |[[ A \land B ]]_W \land [ [ B ] ]_W \cup ( [ [ A \land C ] ]_W \land [ [ C ] ]_W) | \]

In this case, we also adopt the determiner some as a neutral representation.

### 5.3 Contraposition

For all linguistic expressions A, B and C such that 

\[ |A|_W < |C|_W , |B|_W < |C|_W , \text{ then:} \]

by-def A B \implies by-def (C \land \neg B) (C \land \neg A).

**For example:**

(2) Most workers are union members entails:

(3) Most people who are not workers.

If C is the set of all entities of the current world W then contraposition permits to express that:

by-def A B \implies by-def \neg B \land A.  (Extended contraposition).

This property goes beyond a theorem given in [van Eijck 84] which states that:

A quantifier Q observes contraposition iff Q is of the form: 'at most k and A are not B'.

In the sense that (1) k is not implicitly intended to be quite small (with respect to the size of the world) and (2) k can be null. In fact, the value of k turns out to be irrelevant since each time a default rule is used a test of coherence is made on a formula.

### 5.4 Cosymmetry

As shown in [van Eijck 84], if C = B, then we have the property of cosymmetry:

\[ \text{by-def A (A \land \neg B) } \implies \text{by-def B (B \land \neg A)}. \]

Furthermore, since reflexivity holds for all A:

\[ \text{by-def A A} \]

we have:

\[ \text{by-def A B } \iff \text{by-def A (A \land B)}. \]

Suppose that: B = \neg B', then:

\[ \text{by-def A B } \iff \text{by-def A (A \land \neg B')} \]

and, finally, from this result and the definition of cosymmetry, we have:

\[ \text{by-def } \neg B \iff \text{by-def B } \neg A. \]

Notice that, due to symmetry of formulae, there is an equivalence instead of an implication.

**For example:**

(8) *Most teenagers are not married* is logically equivalent to

(9) *Most married people are not teenagers.*

This latter result can also be used to build passive forms from their affirmative counterparts.

### 5.5 Subalternacy

From complex relations holding between different classes of determiners, the property of subalternacy has emerged and turns out to be relevant for statements represented by default logic. This property states that:

For all A, B, by-def A B \implies \neg by-def A (A \land \neg B).

For the same reasons as above, this expression can be simplified and becomes:

by-def A B \implies \neg by-def A \land B.

(10) *Most birds fly* entails

(11) *It is false that most birds cannot fly.*

or, using contraposition and if few is the opposite of most:

(12) *Few birds do not fly.*
6 Epilogue

Default logic permits to represent with a greater accuracy and relevance several types of words and constructions. We have proposed here to integrate it into the formal framework of Generalized Quantifiers. Default logic also exhibits a number of particular properties at the level of the characterization of the truth persistence of a statement represented by default logic in a knowledge base being updated. Next, new and revised inferential patterns are introduced and illustrated. These patterns permit to derive new default rules and to construct new linguistic expressions from previous ones.

The forthcoming works include:

1. The extension and investigation of those linguistic expressions that can be represented by default logic.
2. A formal characterization of contextual situations where default logic can (or must) be used, permitting thus the specification of contextual semantic compositional rules, coherent with the representation defined in [Saint-Dizier 86].
3. The establishment of a link with the logic of presuppositions.
4. The integration of default logic into other formal theories of natural language semantics, in particular into the framework of Situations Semantics [Barwise and Perry 83].

Acknowledgements I am very grateful to Philippe Besnard, Mario Borillo and Jim Delgrande for their useful comments on this work. I thank also Dag Vesterstahl for providing me several of his publications mentioned below. This work was supported by the French INRIA and the PRC CNRS Communication Homme-machine, both civilian public institutions.

References

Barwise, J., Cooper, R., Generalized Quantifiers and Natural Language, *Linguistics and Philosophy*, 4, 159-219, 1981.
Barwise, J., Perry, J., *Situations and Attitudes*, MIT Press, 1983.
Besnard, P., *An Introduction to Default Logic*, Symbolic computation series, Springer-Verlag, 1989.
Van Benthem, J., *Essays in Logical Semantics*, Essays in Linguistics and Philosophy, Vol 29, D. Reidel, 1986.
Delgrande, J., An Approach to Default Reasoning based on a First Order Conditional Logic, *proc. of AAAI-87*, Seattle, WA, 1987.
Dunin-Keplitz, B., Default Reasoning in Anaphora Resolution, *Proceedings of ECAI-84*, Pisa, North Holland, 1984.
Van Eijck, J., Generalized Quantifiers and Traditional Logic, in: *Generalized Quantifiers and Natural Language*, J. Van Benthem and A. Ter Meulen Edts, GRASS series no. 4 pp 1-19, Foris, Dordrecht, 1984.
Joshi, A., Webber, B., Weischedel, R., Some Aspects of Default Reasoning in Interactive Discourse, in: *Communication Failure in Dialogue and Discourse*, R.G. Reilly Edt., North Holland, 1987.
Mercer, R., Using Default Logic to Derive Natural Language Presuppositions, *proceedings of CScS188*, Edmonton, 1988.
Reiter, R., A Logic for Default Reasoning, *Artificial Intelligence*, Vol 13, pp 81-132, 1980.
Saint-Dizier, P., An Approach to Natural Language Seman-