Commonsense Knowledge: Papers from the AAAI Fall Symposium (FS-10-02)

Quantificational Sharpening of Commonsense Knowledge

Jonathan Gordon
Department of Computer Science
University of Rochester
Rochester, NY, USA
jgordon@cs.rochester.edu

Lenhart K. Schubert
Department of Computer Science
University of Rochester
Rochester, NY, USA
schubert@cs.rochester.edu

Abstract

The KNEXT system produces a large volume of factoids from text, expressing possibilistic general claims such as that 'A PERSON MAY HAVE A HEAD' or 'PEOPLE MAY SAY SOMETHING'. We present a rule-based method to sharpen certain classes of factoids into stronger, quantified claims such as 'ALL OR MOST PERSONS HAVE A HEAD' or 'ALL OR MOST PERSONS AT LEAST OCCASIONALLY SAY SOMETHING' – statements strong enough to be used for inference. The judgement of whether and how to sharpen a factoid depends on the semantic categories of the terms involved and the strength of the quantifier depends on how strongly the subject is associated with what is what is predicated of it. We provide an initial assessment of the quality of such automatic strengthening of knowledge and examples of reasoning with multiple sharpened premises.

Introduction: From Weak Knowledge to Strong Knowledge

Human-level artificial intelligence, as required for problems like natural language understanding, seems to depend on the ability to perform commonsense reasoning. This in turn requires the availability of considerable general world knowledge in a form suitable for inference. There are several approaches to acquiring such knowledge, including directly interpreting general statements such as glosses in dictionaries (*e.g.*, Clark, Fellbaum, and Hobbs 2008), abstracting from clusters of propositions (Van Durme, Michalak, and Schubert 2009), and the hand-authoring of rules, as in Cyc (Lenat 1995). Hand-coding is apt to be haphazard in its relationship to language as it depends on the cerebration of numerous knowledge engineers with differing intuitions and no particular commitment to consistency with language or across domains. The volume of hand-coded knowledge produced so far is also probably a couple orders of magnitude short of what is needed.

The approach of this paper is to begin with the large volume of weak, general "factoids" discovered by the KNEXT (KNowledge EXtraction from TExt) system (Schubert 2002), select factoids that lend themselves to logical strengthening, and then sharpen these into quantified general statements that can be used in combination with other facts to generate new conclusions by forward inference.

For example, the following is a factoid and its automatic English verbalization:

[⟨det elm_tree.n⟩ have.v ⟨det branch.n⟩]
'AN ELM TREE MAY HAVE A BRANCH'

Using interpretive rules and abstraction rules, KNEXT obtains such factoids from parsed English text. In this case, the factoids was extracted from some text that referred to a 'branch of an elm tree'. A text occurrence like this indicates that *at least sometimes* an elm tree has a branch, so this is how we understand and verbalize the formula.

However, this factoid is not as strong as we would like it to be: It can be strengthened to say that all – or at least most – elm trees have a (*i.e.*, at least one) branch and that having this branch is an episode that is generally permanent with respect to the tree's existence. We write this using quantifier restrictors as:

(all-or-most *x*: [*x* elm_tree.n]
(some *e*: [[*x* | *e*] permanent]
(some *y*: [*y* branch.n]
[[*x* have-as-part.v *y*] ** *e*])))

The predicate *have* is used in many ways, and we have disambiguated it here to *have-as-part*. Other senses of *have*, such as *give-birth-to* or *possess*, and other kinds of predicates will require different quantification.

In this paper we present a general, rule-based method of doing such quantificational sharpening using existing lexical semantic categories and corpus frequencies. We change unscoped quantifiers to scoped ones and estimate the frequency of events/times and subjects for which each factoid is likely to hold. We then show some simple examples of commonsense reasoning using multiple sharpened premises.

Background and Logical Form

Our perspective on commonsense knowledge derives from our view that the linguistic expression of commonsense knowledge and commonsense reasoning needs to be taken seriously if we wish to build knowledge bases that can support genuine language understanding and reasoning with linguistically derived information. This is why we use a Montague-inspired logic allowing for generalized quantifiers, Episodic Logic (*e.g.*, Schubert and Hwang 2000) as our representation of linguistic meaning and general knowledge rather than, for instance, resorting to nonmonotonic reasoning (NMR).

NMR would capture generalizations such as that "Most (pet) dogs are friendly" through a rule of the type "if *x* is a dog, you can conclude that *x* is friendly, unless you can prove otherwise". However, in general this is neither effective (provability is undecidable even in FOL) nor usable as a premise allowing us to infer, say, that many dogs are friendly (given that there are very many dogs), nor easily adaptable to other nonclassical quantifiers, such as *many* or *occasionally* (*e.g.*, in 'Occasionally, a tree is struck by lightning in a thunderstorm'). But such quantified facts are important in language, commonsense reasoning, and life.

Episodic Logic (EL) uses infix form in square brackets for predication (*i.e.*, with the sentence subject preceding the predicate); *e.g.*, [Romeo.name loves.v Juliet.name], rather than loves.v(Romeo.name, Juliet.name). Note also that individual constants, predicate constants, and most other constants have an extension indicating their NL-derived type. In terms like ⟨det person.n⟩, *det* (for "determiner") would have been something like 'a', 'this', 'many', *etc.* in the original text but has been abstracted to just meaning 'a' or 'an'. The angle brackets indicate an unscoped quantifier, and such terms can be raised to take scope over the sentence they occur in. *E.g.*, [⟨det boy.n⟩ sneeze.v] becomes (det *x* [*x* boy.n] [*x* sneeze.v]) or, using *some* instead of *det*, we have (some *x* [*x* boy.n] [*x* sneeze.v]), where [*x* boy.n] is functioning as a quantifier restrictor. EL syntax allows both restricted and unrestricted quantifications, so another equivalent rendering is (some *x* [[*x* boy.n] and [*x* sneeze.v]]).

Strengthening Factoids

For much of KNEXT's output, the weak formulation is as much as want to assert. So when sharpening, we want to focus on those factoids that are likely targets to be strengthened. The method for doing so is to write rules that match large sets of factoids to patterns using semantic predicates. An example pattern is:

((1_det? 2_animal?) have.v (3_det? 4_animal-part?)),

where the numbers are variables that matching parts of the formula are bound to, and after the underscore is the name of a function that checks whether there is a match. A rule then specifies how these matching terms are to be reordered and transformed to create an appropriately sharpened formula, *e.g.* (in Lisp notation),

(all-or-most *x* (*x* 2_)
  (some *e* ((*x* . *e*) permanent)
    (some *y* (*y* 4_) ((*x* have-as-part.v *y*) ∗∗ *e*))))

Without sharpening, KNEXT learns that 'A PERSON MAY HAVE A HEAD', but we know that having a head isn't optional: it's a crucial part of being a (living) person. Even for body parts that can be lost, it's reasonable to conclude that *most* people have them, so this is what the rule asserts.

To identify an *animal_part* above, we make use of the WordNet (Fellbaum 1998) hierarchy for nouns, which classifies most of these as hyponyms of *body part*. Similar rules can be used to match plants with plant parts and artifacts with their parts. As an additional distinction, we want to know if something is only a ephemeral part, such as a leaf on a tree, in which case it is inappropriate to say that having the part is permanent with respect to the tree's existence. Such cases are few and can be hand enumerated.

Note that the *part-of* relations expressed in sharpened factoids needn't be in WordNet. For example, in a factoid of type [⟨det contraption.n⟩ have.v ⟨det button.n⟩], we would interpret this is involving have-as-part.v as long as WordNet treats *some* sense of 'button' as part of *something* such as a shirt, doorbell, cellphone, *etc.* The same is true for factoids like [⟨det chicken.n⟩ have.v ⟨det feather.n⟩] or [⟨det rosebush.n⟩ have.v ⟨det flower.n⟩].

Disambiguation

The most frequent relation in KNEXT-extracted factoids is 'have', so disambiguation of this "light" predicate, to the extent necessary for inference, is of particular interest. However, not all word senses need to be disambiguated. We claim that verbal predicates aren't nearly as ambiguous as has generally been assumed; they're just semantically *general*.

We take the criterion for whether disambiguation is necessary to be whether or not the entailments follow from the argument types. For example, it's not strictly necessary to disambiguate 'have' in 'have an accident' since the only possible entailments of this phrase in actual use are those for the *experience* sense. By contrast, it is important for us to be able to narrow the sense of 'have' to *eat* in 'A PERSON MAY HAVE A LOBSTER' if that (rather than a possessive sense) is the intended meaning. So the appropriate sharpening would be as follows (where *e* is the eating episode characterized by the sentence, with the characterization relation indicated by the episodic operator '∗∗'):

[⟨det person.n⟩ have.v ⟨det lobster.n⟩]

(many *x*: [*x* person.n]
    (some *e* (some *y*: [*y* lobster.n]
        [[*x* eat.v *y*] ∗∗ *e*])))

Note that *have* often simply serves as a kind of particle binding a *relational* noun to the subject, as when we say 'John has a sister' or 'John has a (certain) weight'. It seems pointless to invent separate meanings of *have* for all these cases, such as *have-as-relative* or *have-as-weight* – these meanings are already inherent in the nominals themselves:

[⟨det male.n⟩ have.v ⟨det sister.n⟩]

(many *x*: [*x* male.n]
    (some *e* (some *y*: [*y* female.n]
        [[*x* (has-as sister.n) *y*] ∗∗ *e*])))

[⟨det male.n⟩ have.v ⟨det weight.n⟩]

(many *x*: [*x* male.n]
    (some *e*: [[*x* | *e*] permanent]
        (some *y* [[*y* weight-of.n *x*] ∗∗ *e*])))

These relational uses of 'have' are identified based on the semantic categories of the subject (*e.g.*, a *causal agent* or *social group*) and the object (*e.g.*, a hyponym of *relative*, *leader*, or *professional*) while most features like 'weight' are hyponyms of *attribute*.

In some cases, disambiguation is necessary but is difficult enough that we choose not to sharpen the factoid rather than risk doing so incorrectly. A particularly difficult class of factoids to sharpen are those involving prepositions, where we need to at least implicitly disambiguate different uses. For instance, 'a man with one arm' is an individual-level use while 'a man with a cake' is stage-level. To avoid bad sharpened output, we also need to check for nouns that don't mean much when standing alone, *e.g.*, 'front', 'thing', or 'issue'. We want to avoid sharpening factoids involving such terms, at least when they occurr as the subject of sentences with no object.

Individual and Stage-Level Predicates

A key distinction we need when strengthening is between individual-level and stage-level predicates (Carlson 1977; Kratzer 1995). Individual-level predicates endure over most of the existence of the individual they're predicated of while stage-level predicates describe dynamic goings-on or transient situations. So while we want to quantify stage-level predicates over individuals and episodes, an individual-level predicate is just quantified over individuals.

We assume that if an entity has a capacity, it is at least occasionally exercised. Thus we sharpen factoids about abilities to stage-level quantification over episodes of performing them. KNEXT generates some factoids about abilities that explicitly state that an individual may be *able* to do something:

[⟨det female.n⟩ able.a (Ka speak.v)]
'A FEMALE CAN BE ABLE TO SPEAK'

(The Ka operator indicates a kind of action.) Factoids like this can indicate abilities that are specific to a few individuals – say, being able to ride a horse – rather than generally true as in the example above. But they can also indicate basic abilities: We rarely state that someone is 'able to' do a basic action like walking. Yet, if someone breaks their leg, we might say that they are 'able to walk (again)' and can produce an appropriate factoid.

Sharpened factoids about abilities are also formed from factoids about actions without *able.a*, such as [⟨det female.n⟩ swim.v]. As a stage-level predicate, *swim* will lead to quantification over episodes:

(many *x*: [*x* female.n]
(occasional *e* [[*x* swim.v] ∗∗ *e*]))

What we aim to get are formal versions of habituals like

All or most people occasionally use a cellphone.
All or most companies occasionally announce a product.

rendered in the following manner:

(all-or-most *x*: [*x* person.n]
(occasional *e* (some *y*: [*y* cell_phone.n]
[[*x* use.v *y*] ∗∗ *e*])))

(all-or-most *x*: [*x* company.n]
(occasional *e* (some *y*: [*y* product.n]
[[*x* announce.v *y*] ∗∗ *e*])))

Stage-level adjectives also get quantified over episodes:

[⟨det male.n⟩ hungry.a]

(all-or-most *x*: [*x* male.n]
(occasional *e* [[*x* hungry.a] ∗∗ *e*]))

Some 'have' propositions represent temporally quantified occurrences, *e.g.*, 'All or most persons occasionally have a thought/cold/shock/party…' We recognize such a use by a subject who is a *causal agent* and an object that is a *psychological feature*, *event*, or *state*.

[⟨det male.n⟩ have.v ⟨det thought.n⟩]

(all-or-most *x*: [*x* male.n]
(occasional *e* (some *y*: [*y* thought.n]
[[*x* experience.v *y*] ∗∗ *e*])))

It would be distressing if we quantified the stage-level verb *die* similarly. For this reason, stage-level predicates are divided into repeatable and non-repeatable ones. The latter includes strict once-per-existence predicates like *die* and also "pivotal" ones like *marry*. While marriage is repeatable, we don't want to claim it's a *frequent* action for an individual. It is also necessary to distinguish those predicates that are nonrepeatable with respect to their objects, *e.g.*, while one can kill multiple times, one can only be killed once. Nonrepeatable predicates generally fall into a small number of VerbNet (Kipper-Schuler 2006) categories, which we supplement with the other terms from the corresponding WordNet synsets.

One class of factoids that should be strengthened to individual-level quantification are nouns predicated of noun phrases, such as [⟨det person.n⟩ ⟨fighter.n⟩]. For these, the noun predicate in object position is likely true of the individual's whole existence, so we sharpen it as

(many-or-some *x*: [*x* person.n]
(some *e*: [[*x* | *e*] permanent]
[[*x* fighter.n] ∗∗ *e*]))

Copular *be* with a noun complement can be treated similarly. We avoid quantifying very strongly when the complement is a role noun, such as those found in WordNet as hyponyms of *person* and ones like *weapon*: While it is true that a person can be a doctor, we don't want to conclude *most* people are doctors. And while a cane can be weapon (when it's used as such), we don't want to assert that many canes *are* weapons.

Most stative verbs are individual-level, but there are exceptions such as *own* or *contain*. This can depend on the subject: Books contain information for the duration of their existence, but a jug only contains water at specific times. Many people *at some time* own a house, but it is not necessary that they own it for their whole life.

We assume that a factoid expresses possession if the verb is *have* or *own*, there is an animate subject (such as a person or organization – hyponyms of *causal agent*), and the object is an artifact or a domesticated animal:

[⟨det person.n⟩ have.v ⟨det dog.n⟩]

(many *x*: [*x* person.n]
(some *e* (some *y*: [*y* dog.n]
[[*x* possess.v *y*] ∗∗ *e*])))

This also illustrates the utility of corpus frequencies: We want to conclude that most men have shoes, but few men have a yacht. We use the strength of the association between the subject of the factoid and what is predicated of it, measured as the pointwise mutual information, as an indicator of how strongly the sharpened formula should be quantified. Taking our formulas as formal generic statements, this approach reflects the inductive view of generalizations: After we observe enough people possessing dogs (from textual references to it), we take it to be likely. For discussion of whether generic statements (such as the commonsense knowledge we are trying to abstract from possibilistic factoids) should be understood inductively or from a rules-and-regulations view not dependent on real-world activity, see Carlson (1995).

Note that many verbs have both stative and dynamic senses, *e.g.*, *think* is stative in 'I think she's tired' but not in 'Think about the problem'. In sharpening factoids we don't generally attempt to distinguish between such uses – we allow a verb to be strengthened as a stative (and thus probably individual-level) if it has any stative sense.

Individual-level adjectives can be found by looking at the hypernyms of the derivationally related form in WordNet, so for *fond* we get *fondness*, which has *attribute* as a hypernym. This, *tendency*, and *quality* are good indicators of an individual-level property, *e.g.*,

[(K (plur cat.n)) fond.a (of.p (K milk.n))]

(most *x*: [*x* cat.n]
[*x* fond.a (of.p (K milk.n))])

The above is a factoid about kinds (indicated by the K operator), which we sharpen to be about individuals of the kind. An additional type of factoid we haven't dealt with here is the type involving kind-level predicates, which are predicated not of individuals but of a whole genus, *e.g.*,

[(K (plur cow.n)) widespread.a]

A problem here concerns the level of abstraction: If we view this factoid as a statement about an individual, *viz.*, the kind *cows* (much as in 'The Milky Way is widespread'), we should not read it possibilistically as 'Cows may be widespread', but simply as 'Cows are widespread'. But when we abstract from the particular kind to species, we want to conclude that *some species* are widespread.

Another special case are factoids with event nominal (*e.g.*, a war or a party) subjects, for which neither stage-level nor individual-level predicates should result in quantification over episodes. We identify these event nominals by using the verbalizations in the Nomlex nominalization lexicon and those words categorized as hyponyms of *event* and related synsets in WordNet.

It's worth noting that while we use WordNet as our primary resource for semantically categorizing predicates in the process of sharpening, our factoids express information beyond what's in WN: While it tells us that *writing* is a *human activity*, it does not tell us that people write *letters* and so on. It is only the combination of KNEXT factoids with WordNet, VerbNet, and other resources of lexical semantic features that provides the bulk of the sharpened output.

Evaluation of Sharpening

To evaluate our sharpening methods, we first took a set of propositions extracted from the British National Corpus that were previously evaluated by crowdsourcing on Mechanical Turk (for details, see Gordon, Van Durme, and Schubert 2010). Non-experts were shown the English verbalizations (*e.g.*, 'A MAN MAY HAVE A HEAD') of factoids and asked to rate how well they conveyed accurate commonsense knowledge according to the instructions in Figure 1.

> The statement above is a reasonably clear, entirely plausible, generic claim and seems neither too specific nor too general or vague to be useful:
>
> 1. I agree.
> 2. I lean towards agreement.
> 3. I'm not sure.
> 4. I lean towards disagreement.
> 5. I disagree.

Figure 1: Instructions for judging of unsharpened factoids.

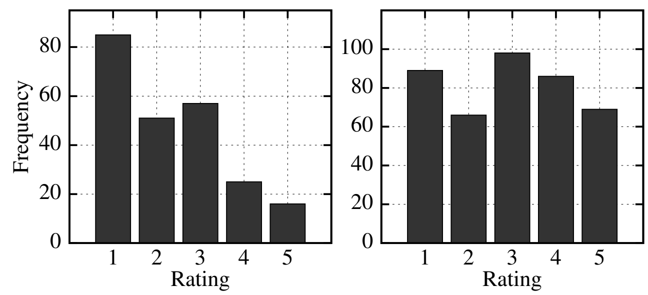

Figure 2: Agreement with Statement 1 for the factoids with unsharpened forms rated 1–2 (left) and all factoids (right). The vertical axis shows the number of factoids that were given each rating, counting both judges' responses.

Out of 1500 randomly sampled BNC factoids, 435 of them could be sharpened. The smaller size of this set represents a preference for precision over recall and the large number of factoids that don't seem to merit a stronger form, even among those that were judged to hold in when stated weakly.

Here we want to judge whether the sharpened forms express reasonable general claims and have been strengthened sufficiently. The authors therefore judged 200 sharpened factoids on the same scale of 1–5 (with 1 being agreement and 5 disagreement) based on their agreement with the following primary and secondary statements:

Statement 1. The factoid is a reasonable general claim about the world even if – perhaps – it isn't as strongly quantified as it might be.

If so (that is, if the judge rates the factoid 1 or 2), they then judged

Statement 2. The quantifiers seem sufficiently strong.

So, for instance,

(some *x*: (male.n *x*)
(some *e*: [[*x* | *e*] permanent]
(some *y*: [*y* head.n *y*]
[[*x* have-as-part.v *y*] ** *e*])))

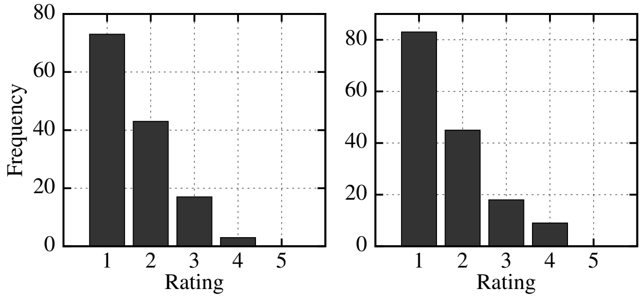

Figure 3: Agreement with Statement 2 for the factoids with unsharpened forms rated 1–2 (left) and all factoids (right)

	Stmt 1 Avg	*Stmt 2 Avg*
Judge 1	3.01	1.66
Judge 2	2.73	1.71
Correlation	0.79	0.75

Table 1: The judges' ratings of the sharpened factoids, including those produced from all unsharpened factoids, not just the highly subset.

would not be rated very well for Statement 2. It is true, but the claim should be quantified more strongly: All men have heads.

Since it is quite hard to produce a good sharpened statement from a bad factoid, we are interested not just in the overall performance of the sharpening but also in how it does on a subset of good factoids. For this, we took those factoids with an average Turker-assigned rating between 1 and 2. Such a rating means that, in its weak, possibilistic form, the factoid is probably a reasonable claim about the world.

As seen in Figure 2 (left), these favorably judged weak factoids yield favorably judged strengthened factoids (when they yield any at all) more often than they yield ambivalent or negative judgements. While 36% of the ratings of factoids sharpened from the good unsharpened factoids (those with an average rating between 1 and 2) are rated a 1, only 21% of the complete set were so rated. As can be seen from the right histogram, judgements of sharpened factoids are considerably worse if the unsharpened factoids include everything generated by KNEXT. Therefore it will be crucial to pre-filter unsharpened factoids, perhaps using crowdsourcing (as was done here) or by improved automatic methods. This can also include improvements in the initial knowledge extraction and in the technology it relies on. Incorrect syntactic parses, including improbable parts-of-speech, were evident in the judged results: Any improvements in parsing are likely to also improve our knowledge extraction.

Inference

From a ground fact like [John.name person.n], we'd like to generate inferences like

(probably
(some *x*: [*x* head.n]
[John.name have-as-part.v *x*]))

(probably
(occasional *e* (some *y*: [*y* cell_phone.n]
[[John.name use.v *y*] ** *e*])))

Given a hierarchy we'd like to be able to generate the above inferences even if we're told just the more-specific knowledge that [John.name male.n]. The reason for the *probably* is that a completely confident answer is not justified by the *all-or-most* quantifier.

Note that we make the slightly unusual assumption that quantifiers like *all-or-most*, *all*, and *most* have "existential import", *i.e.*, unlike in FOL, the quantifiers imply *some*. We even make the stronger assumption that these quantifiers imply *many*. For example, if we are given that all-or-most elm trees have branches, we'll conclude that *many* elm trees have branches. (This simply is an assumption that all the types that occur in our formulas have numerous instances, whether they be people, dogs, times, legs, *etc.*)

Furthermore, we want to sanction some rough-and-ready inferences involving quantifier chaining. For example, using premises 'ALL GOLDFISH ARE PETS', and 'ALL-OR-MOST PETS ARE HARMLESS', we want to conclude, tentatively, that 'ALL-OR-MOST GOLDFISH ARE HARMLESS'. This is expected to hold in the absence of further information, *i.e.*, when we have no reason to suppose that goldfish are exceptional as pets with respect to the 'harmless' property. In the following, since we intended *all-or-most* to suggest somewhat higher proportions than *most*, we will underscore that distinction by rewriting *all-or-most* as *virtually-all*. Formally,

(all *x*: [*x* goldfish.n] [*x* pet.n]),
(virtually-all *x*: [*x* pet.n] [*x* harmless.a])

(virtually-all *x*: [*x* goldfish.n] [*x* harmless.a]))

Note that in such inferences, the quantifier in the conclusion will in general be weaker than the quantifiers in the premises (except that *all* maintains the full strength of whatever the other quantifier expresses). Table 2 shows how these quantifiers combine (under the sorts of *ceteris paribus* assumptions mentioned above: a subset S of a set S′ shows the same distribution of properties as S′, in the absence of information to the contrary). This table is intended to be "not too unreasonable" in relation to ordinary use of these quantifiers, rather than reflecting formal, precise meanings. It has been constructed by associating the following fractional (or probabilistic) lower bounds with (all but a few of) the quantifiers:

	vall	most	vmany	many	qmany	fmany
At least	85%	72%	60%	50%	35%	25%

Table 3: Lower bounds for quantifiers.

Note that these are being interpreted as *proportional* quantifiers, not absolute ones. The quantifier *all* of course is also proportional, and means 100%. By multiplying these when we chain two inferences, we get the above results.

Some is not being treated as proportional – it just means "at least one". We haven't shown *occasional*, because this

Q1 Q2	*all*	*vall*	*most*	*vmany*	*many*	*qmany*	*fmany*	*some*
all	all	vall	most	vmany	many	qmany	fmany	some
vall	vall	most	vmany	many	qmany	fmany	some	some
most	most	vmany	many	qmany	qmany	fmany	some	some
vmany	vmany	many	qmany	qmany	fmany	some	some	some
many	many	qmany	qmany	fmany	fmany	some	some	some
qmany	qmany	fmany	fmany	some	some	some	some	some
fmany	fmany	some	some	some	some	some	some	some
some	some	some	some	some	some	some	some	some
few	few	few	?	?	?	?	?	?
no	no	few	?	?	?	?	?	?

Table 2: A table of quantifier combinations, writing *vall* for *virtually-all* and adding *vmany* for *very-many*, *qmany* for *quite-many*, *fmany* for *fairly-many*, and *few*. Q1 is the first quantifier in the chain and Q2 the second one.

isn't proportional either, and should be used only for temporal (or event) quantification; but as such, it is stronger than *some* – for example, if Kevin occasionally smokes, and when he smokes he occasionally coughs, then he occasionally (not just at least once!) coughs, even though at a lower frequency than he smokes.

Few (in many uses) can be regarded as proportional, meaning something like $<15\%$. But chaining it with other quantifiers is a subtle matter. For example, does "Few P are Q and few Q are R" imply "Few P are R"? No; *e.g.*, the following is grossly unsound: "Few people are Quakers and few Quakers are Republicans, therefore few people are Republicans". The same goes for quantifier *no*. (* "No dogs are chickens and no chickens have four legs, therefore no dogs have four legs"). In fact, with *few* or *no* as first quantifier (Q1 in the table), no meaningful chaining seems to be possible, so they are only shown as second quantifier (Q2).

Conclusions

In this paper we have suggested that we can sharpen classes of factoids expressing commonsense knowledge from weak, possibilistic claims to stronger, quantified claims. To do this we made use of semantic categories for disambiguating predicates and recognizing factoids that express characterizing properties that deserve strengthening. We also made use of the frequency with which we extract claims from text to induce the strength of that quantification. Initial evaluation suggests that the resulting strengthened factoids are of good quality, though improvement is needed for them to be suitable for inference. We've also suggested how such generalized quantifiers could be combined when performing inference, which makes this kind of knowledge desirable for enabling commonsense reasoning.

Acknowledgements

This work was supported by NSF grants IIS-0535105 and IIS-0916599 and by subcontract FY10-01-0297 under ONR STTR N00014-10-M-0297.

References

Carlson, G. N. 1977. *Reference to Kinds in English*. Ph.D. in Linguistics, University of Massachussetts, Amherst. Published 1980 by Garland Press, New York.

Carlson, G. N. 1995. Truth-conditions of Generic Sentences: Two Contrasting Views. In Carlson, G. N., and Pelletier, F. J., eds., *The Generic Book*. University of Chicago Press.

Clark, P.; Fellbaum, C.; and Hobbs, J. 2008. Using and Extending WordNet to Support Question-answering. In *Proceedings of the 4th Global WordNet Conference (GWC '08)*.

Fellbaum, C. 1998. *WordNet: An Electronic Lexical Database*. MIT Press.

Gordon, J.; Van Durme, B.; and Schubert, L. K. 2010. Evaluation of Commonsense Knowledge with Mechanical Turk. In *Proceedings of the NAACL 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk*.

Kipper-Schuler, K. 2006. *VerbNet: A Broad-Coverage, Comprehensive Verb Lexicon*. Ph.D. Dissertation, University of Pennsylvania.

Kratzer, A. 1995. Stage-level/individual-level Predicates. In Carlson, G. N., and Pelletier, F. J., eds., *The Generic Book*. University of Chicago Press.

Lenat, D. B. 1995. Cyc: A Large-scale Investment in Knowledge Infrastructure. *Communications of the Association for Computing Machinery* 38(11):33–48.

Schubert, L. K., and Hwang, C. H. 2000. Episodic Logic Meets Little Red Riding Hood: A comprehensive, natural representation for language understanding. In Iwanska, L., and Shapiro, S. C., eds., *Natural Language Processing and Knowledge Representation: Language for Knowledge and Knowledge for Language*. MIT/AAAI Press.

Schubert, L. K. 2002. Can We Derive General World Knowledge from Texts? In *Proceedings of the 2nd International Conference on Human Language Technology Research (HLT02)*.

Van Durme, B.; Michalak, P.; and Schubert, L. K. 2009. Deriving Generalized Knowledge from Corpora using WordNet Abstraction. In *Proceedings of EACL*.