Automatically Constructing Concept Hierarchies of Health-Related Human Goals

Mark Kröll[1], Yusuke Fukazawa[2], Jun Ota[3], and Markus Strohmaier[1]

[1] Graz University of Technology and Know-Center
Inffeldgasse 21a, 8010 Graz, Austria
{mkroell,markus.strohmaier}@tugraz.at
[2] NTT DOCOMO, Inc.
3-6, Hikari-no-oka, Yokosuka, Kanagawa, Japan
{fukazawayuu}@nttdocomo.co.jp
[3] The University of Tokyo
5-1-5 Kashiwanoha, Kashiwa, Chiba, Japan
{ota}@race.u-tokyo.ac.jp

Abstract. To realize the vision of intelligent agents on the web, agents need to be capable of understanding people's behavior. Such an understanding would enable them to better predict and support human activities on the web. If agents had access to knowledge about human goals, they could, for instance, recognize people's goals from their actions or reason about people's goals. In this work, we study to what extent it is feasible to automatically construct concept hierarchies of domain-specific human goals. This process consists of the following two steps: (1) extracting human goal instances from a search query log and (2) inferring hierarchical structures by applying clustering techniques. To compare resulting concept hierarchies, we manually construct a golden standard and calculate taxonomic overlaps. In our experiments, we achieve taxonomic overlaps of up to ~51% for the health domain and up to ~60% for individual health subdomains. In an illustration scenario, we provide a prototypical implementation to automatically complement goal concept hierarchies by means-ends relations, i.e. relating goals to actions which potentially contribute to their accomplishment.

Our findings are particularly relevant for knowledge engineers interested in (i) acquiring knowledge about human goals as well as (ii) automating the process of constructing goal concept hierarchies.

Keywords: Knowledge acquisition, human goal knowledge, goal concept hierarchy, means-ends relation.

1 Introduction

A better understanding of what motivates humans to perform certain actions is relevant for a range of challenging research problems. These problems include goal recognition from people's actions, reasoning about people's goals or the generation of action sequences, i.e. planning [1]. Reasoning, for instance, helps to answer *why* questions and can thus support intelligent agents in their decision making processes.

H. Xiong and W.B. Lee (Eds.): KSEM 2011, LNAI 7091, pp. 124–135, 2011.

To enable planning or reasoning, knowledge about human goals needs to be structured and organized, e.g. by arranging it in hierarchical structures. In this context, hierarchies of goal concepts have proven valuable in several research areas including (i) web search ([2], [3]), (ii) intelligent user interfaces ([4], [5]) or (iii) semantic task retrieval ([6], [7]). Concept hierarchies are meant to mimic mental constructs thus reflecting a domain's abstract representation. (i) In web search, goal concept hierarchies organize users' underlying search goals to inform and improve search engines' retrieval performances. (ii) By utilizing structured human goal knowledge, intelligent user interfaces are capable of better understanding relationships between people's goals and their actions. (iii) With respect to semantic task retrieval, finding appropriate web services is facilitated by a better understanding of which tasks or actions are required to accomplish people's goals. In [6], we have already experimented with modelling goal structures and connecting them to web services.

In this work, we seek (i) to automatically construct a hierarchy of goal concepts and (ii) to study the extent to which automating this process is feasible. In our experiments, we organize health-related human goal concepts in a hierarchy, thus explicating goal knowledge from the health domain. The construction process consists of following two steps:

1. **Extracting health-related human goals from a search query log:** We extract a set of health-related human goal instances from the AOL search query log [8]. We then apply techniques from previous work [9] to obtain a set of ~500 health-related human goal concepts. To evaluate automatic construction approaches, two of the authors manually arrange these concepts into a hierarchy and thereby craft a golden standard (see Section 2).
2. **Automatically inferring hierarchical structures:** To hierarchically relate goal concepts, we explore the potential of three established techniques, i.e. Bi-Section K-Means, Formal Concept Analysis and Hierarchical Agglomerative Clustering. These techniques were successfully applied in the past to construct concept hierarchies (cf. [10]). To be able to compare these techniques in terms of quality, we calculate taxonomic overlaps [11] between resulting concept hierarchies and the golden standard.

Automating the construction of goal concept hierarchies addresses the goal acquisition bottleneck [12] which refers to the costs associated with the acquisition process. In Section 3, we explore potentials of algorithmic approaches and achieve taxonomic overlaps of up to ~51% for the entire domain and up to ~60% for individual health subdomains.

Section 4 presents an illustration scenario where we seek to automatically complement goal concept hierarchies with means-ends relations, i.e. relating goals to actions which contribute to their accomplishment. To give an example, the action "use condoms" potentially contributes to accomplish the health-related goal "prevent aids". Enriching goal concept hierarchies by means-ends relations enables more complex operations on goal knowledge, e.g. the generation of action sequences in planning procedures.

This work's findings are relevant for knowledge engineers interested in (i) acquiring knowledge about human goals as well as (ii) automating the process of constructing goal concept hierarchies.

2 Health-Related Human Goals

In this section, we describe how we tap into search query logs as a source for extracting (health-related) human goal instances (cf. [9]). We then transform these instances into concepts and manually organize them in a hierarchy. This hand-crafted hierarchy will serve as golden standard in our experiments.

2.1 Health Related Goals from Search Query Logs

Search query logs are a particularly valuable source for extracting human goal instances. Each submitted query expresses a user's search goal expressed either explicitly or implicitly. In the following, we present some examples of health-related queries which do or do not contain explicit human goal instances (obtained from [8]):

Queries containing explicit goal instances	Queries not containing explicit goal instances
"lose weight fast"	"weight loss"
"writing medical case studies"	"case study research"
"passing a drug test"	"drug test"

In [9], we developed an algorithm to identify those queries which contain explicit goal instances. Our algorithm automatically extracts ~90.000 queries of which 77 out of 100 queries actually contained explicit goal instances (77% precision). We then filtered the ~90.000 queries with regard to (i) health-related keywords such as "healthy" or "disease" and (ii) health-related URLs such as "http://www.camh.net/" or "http://www.healthandage.com/". Keywords as well as URLs were compiled and refined manually, i.e. by means of brainstorming sessions, manual inspection and using the open directory project[1]. The keyword-based approach identified fewer health-related goal queries than the URL-based approach, yet with a higher accuracy, i.e. 73.2% over 44.8%. To gather a useful set of health-related goal queries, we combined both filtering approaches. We then removed duplicates and false positives such as "follow your heart", "donate your car" or "find healthy dog food".

To reduce ambiguity, we conceptualize these health-related goal instances, i.e. we convert them into goal concepts. We start forming concepts by normalizing individual instances of human goals, i.e. by removing stop words and punctuations, by lemmatizing verbs, i.e. reducing them to their base form and by transforming nouns to their singular form like similar work by [13]. A concept then encompasses all instances that normalize to the same text as illustrated in the following:

Human Goal Concepts	Corresponding Human Goal Instances
"increase health"	"increasing health", "increase your health"
"lose weight"	"lose some weights", "losing a lot of weight"

Consequently, goal concepts have a correspondence to verb phrases. Yet, goal concepts are not literal strings of text but stand for mental artifacts. A concept can represent related or synonymous instances of human goals.

[1] http://www.dmoz.org/

For our experiments, the set of health-related goal concepts totals 489. We deem this set size sufficient for our purposes since our work focuses on exploring means to automatically construct goal concept hierarchies. Large-scale experiments appear to be warranted whence adequate techniques have been identified.

2.2 Golden Standard

To the best of our knowledge, there is no publicly available golden standard which hierarchically relates health-related goal concepts. To that end, we decided to handcraft a golden standard which would enable us to compare algorithms and thus metrically assess the quality of resulting concept hierarchies. The generation of the golden standard consists of three steps: (1) to group ~500 health-related goal concepts by similar topics, e.g. pregnancy-related goal concepts, (2) to hierarchically relate these topic groups, e.g. to connect plastic surgery-related and skin-related goals by introducing the higher level cluster beauty-related goals, and eventually (3) to hierarchically relate goal concepts within each group. Table 1 shows the top-level categories of the golden standard after the first two steps.

Table 1. Shows the golden standard's 1st and 2nd level categories

1st Level Categories	2nd Level Categories
Sex & Baby	Sex, Pregnancy, Baby
Health & Beauty	Daily Health-Care, Healthy Diet, Weight-Control, Fitness, Beauty, Face-Care, Mental-Health
Disease	Body-Disease, Brain-Disease
Other	Authority, Drugs

In the third step, we introduce hierarchical structures within each topic group. For each group, we select general goal concepts as higher level node candidates. General goal concepts tend to consist of a verb and only few objects. Then, we look for specializations of general goal concepts. We consider two kinds of specializations: 1) object specialization, e.g. "buy cheap diet pill" is an object specialization of "buy diet pill" and 2) method specialization, e.g., "prepare weight loss diet" is a method specialization of "lose weight". These two steps are repeated until no other concepts remain in the respective group. Fig. 1 visualizes a small excerpt from the resulting golden standard.

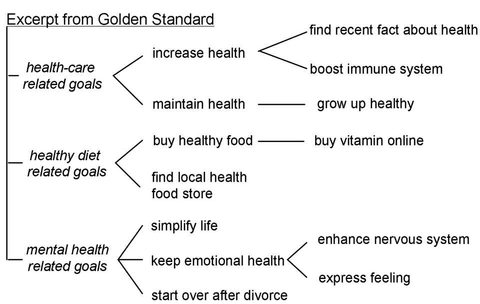

Fig. 1. Visualizes a small excerpt from the golden standard. Higher-level categories are inserted by hand when no appropriate goal concept was present (*in italics*).

We conclude this section by analyzing the resulting structure of the golden standard and its components. This analysis might give some indication of health-related concerns as well as e-health trends (at that time). The golden standard's four top level groups partly reflect people's health conditions, their situations in life as well as what their fears and problems are. The first group *Sex & Baby* comprises all goal concepts related to sexual activities as well as consequences, e.g. pregnancy, baby care, yet also dealing with AIDS. The second topic group *Health & Beauty*, on the one hand, reflects people's longing for an immaculate appearance, e.g. by having plastic surgery. On the other hand, it reveals a trend towards a healthier lifestyle including diet, fitness and weight issues. The third group *Disease* contains goal concepts which are related to physical dysfunctions, e.g. "eliminate large kidney stone" or "go off seizure medication". The fourth group *Other* comprises all remaining health-related goal concepts which are further divided into *authority-related* and *drugs-related* goals.

At this point, we mention the presence of goals which people did not want to achieve but rather avoid. To give some examples, selected avoid goals include: "keep getting sick", "have nocturnal seizure", "have dark circle under eye", "lose mind", "lose hair" or "catch aids". While we did not include these goals in our golden standard, we deem them interesting for they provide us with insights about people's fears and concerns.

3 Goal Concept Hierarchy

In this section, we explore the potential of three algorithmic approaches to automatically infer concept hierarchies of health-related human goals. To compare these approaches, we calculate taxonomic overlaps between resulting concept hierarchies and our golden standard.

3.1 Experimental Setup

We examine three algorithms which are introduced in [14] as viable practices to automatically construct concept hierarchies: Formal Concept Analysis (FCA), Hierarchical Agglomerative Clustering (HAC), and Bi-Section K-Means (BSKM). We experiment with three feature types: token-based (T), neighborhood-based (N) and click-through-based (C).

Feature Type	Description
token-based (T)	For every goal concept, all corresponding goal instances are tokenized and sanitized, e.g. stop words are removed. Then a characteristic token vector is formed to represent the goal concept.
neighborhood-based (N)	Since we tapped into search query logs to extract human goal instances, we have access to neighborhood information. We assume that neighboring queries are suitable features to represent goal queries. The set of neighboring queries encompasses queries issued by the same user before and after the goal query. To give an example, the query "lose weight fast" possesses a number of neighboring queries including "weight loss supplements", "types of diet pills" or "Lipo6". After tokenization and sanitization steps (as in the token-based approach), we generated a characteristic term vector using a set of neighboring queries without including the goal query.
click-through-based (C)	The search query log also contains click-through information, i.e. given a query which resulting URLs were clicked. For each human goal query, we traverse the search query log and collect all corresponding clicked URLs. To be added to the feature vector, each URL must have been clicked at least twice.

In the following, we provide implementation details for FCA, HAC and BSKM:

Formal Concept Analysis (FCA) originated as data analysis technique before it was successfully applied to construct concept hierarchies [10]. In our experiments, we use an existing java library which provides us with an efficient implementation[2].

Hierarchical Agglomerative Clustering (HAC) is a similarity-based bottom-up clustering algorithm. For our experiments, we implement the HAC algorithm according to [14]. As similarity metric, we choose single linkage since it possesses the lowest computational complexity, i.e. $O(n^2)$, compared to other linkage metrics.

Bi-Section K-Means (BSKM) represents a clustering technique that repeatedly applies the traditional K-Means algorithm. We implement the Bi-Section K-Means algorithm according to [14]. As similarity measure, we use the cosine measure.

As evaluation metric, we utilize the taxonomic overlap which was pioneered by [11] as one of the first metrics to compare two concept hierarchies with each other. The metric allows a comparison not only on a lexical level but also on a conceptual one. The principal idea behind this metric is that two concept hierarchies are similar (i) if they have a lot of concepts in common and (ii) if these common concepts share many super/sub concepts. A high overlap between an automatically constructed concept hierarchy and the hand-crafted golden standard would thus indicate a good quality. To calculate the taxonomic overlap (TO) in respect to one common concept (c), we use following formula:

$$TO(c, O_1, O_2) = \max_{c' \in C_2} \frac{|SC(c, O_1) \cap SC(c', O_2)|}{|SC(c, O_1) \cup SC(c', O_2)|}$$

where O_1 represents the golden standard, O_2 the automatically constructed concept hierarchy, C_2 the set of O_2's concepts and SC stands for semantic cotopy, i.e. the set of concept c's super and sub concepts. For calculating the SC, we do not take into account concepts without name which have been created during the clustering process[3]. Repeating this calculation for all concepts, we obtain an averaged TO which reflects the similarity between two concept hierarchies O_1 and O_2:

$$\overline{TO}(O_1, O_2) = \frac{1}{|C_1|} \sum_{c \in C_1} TO(c, O_1, O_2)$$

where $|C_1|$ stands for the number of concepts in the golden standard O_1. The taxonomic overlap can be calculated in both directions, i.e. $TO(O_1,O_2)$ equaling precision and $TO(O_2,O_1)$ equaling recall. In this work, we seek to explore the potential of three algorithms to hierarchically structure a fixed set of goal concepts; thus, all reported taxonomic overlaps represent precision results.

---

2 http://www.st.cs.uni-saarland.de/~lindig/#colibri by D. Gotzmann

3 Clustering algorithms introduce concepts without name into resulting hierarchies. Yet, concepts without name do not exist in the golden standard. For calculating the SC, we thus do not take these concepts into account to allow for fair comparisons between the clustering algorithms.

3.2 Evaluation

Table 2 summarizes taxonomic overlap results for all (algorithm, feature type) combinations. Each overlap value reflects the degree of matching between an automatically constructed goal concept hierarchy and the golden standard.

Table 2. Shows taxonomic overlaps for (algorithm, feature type) combinations

	Token-Based (T)	Neighborhood (N)	Click-Through (C)
BSKM	48.06 %	45.04 %	39.69 %
FCA	41.52 %	40.91 %	39.39 %
HAC	*50.82 %*	47.55 %	46.11 %

The combination (T, HAC) yields the highest taxonomic overlap value 50.82% and (C, FCA) the lowest value 39.39%. HAC and BSKM appear to be equally well suited for the construction process. All clustering algorithms achieve highest results when using the token-based features. In addition, we experimented with combinations of feature types as well which did not yield better results.

To further investigate the applicability of the clustering algorithms, we calculate taxonomic overlaps for 14 subdomains which differ, e.g. in their number of concepts. The subdomains are formed by taking second level categories (see Table 1) as roots.

Table 3. Shows individual taxonomic overlaps for 14 subdomains (token-based features)

Health Sub Domain	#Concepts	Taxonomic Overlap		
		FCA	BSKM	HAC
Body-Disease	79	32.35%	44.46%	*44.64%*
Daily Health-Care	78	36.26%	41.91%	*50.92%*
Sex	58	35.82%	*47.81%*	46.09%
Beauty	54	38.65%	46.48%	*49.93%*
Weight control	52	32.00%	46.45%	*46.89%*
Pregnancy	51	38.74%	47.16%	*51.36%*
Mental Health	43	40.03%	47.13%	*50.44%*
Face-Care	40	42.98%	48.28%	*49.56%*
Healthy diet	39	39.24%	49.62%	*55.23%*
Fitness	28	41.30%	37.04%	*42.14%*
Baby	25	42.39%	*53.63%*	52.08%
Authority	23	51.13%	54.30%	*59.67%*
Brain-Disease	18	52.43%	49.8%	*54.59%*
Drug	10	58.00%	*58.72%*	57.91%

Results in Table 3 show that the HAC algorithm yields highest taxonomic overlap values in most cases, e.g. 59.67% for the authority subdomain. These results indicate that the degree of overlap does not depend on the number of concepts, i.e. high overlaps were achieved for subdomains with few and many concepts. To learn more about what causes low taxonomic overlaps, we examine the *Fitness* health subdomain which yields a taxonomic overlap of only 42.14%. In Fig. 2, we compare excerpts from the clustering result (T, HAC) to the golden standard.

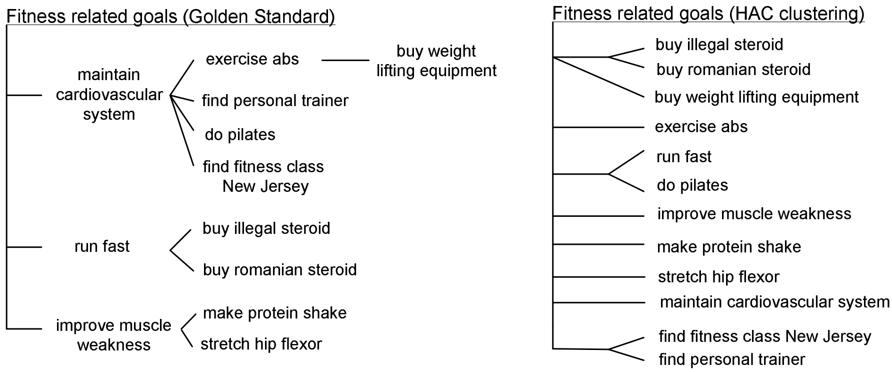

Fig. 2. An excerpt of the golden standard's structure is compared to HAC's respective clustering result. The excerpts show goal concepts from the health subdomain *Fitness*.

From visual inspection, a primary reason for the low overlap appears to be the flattening of the hierarchical structure. This might be a consequence of our feature representations not capturing the required information for generating a hierarchy according to the golden standard. We thus propose to use context information[4] to represent goal concepts. Context information can, for instance, be acquired from social media corpora such as weblogs. This kind of representation might also better comply with FCA which did not seem to reach its full potential with the used features types. Lastly, we can observe an implication of token-based representations: human goal concepts are clustered together based on equal verbs, e.g. "buy" or "find", although they do not belong together from a semantic point of view. To alleviate this issue, we could separate verb from noun semantics and treating, e.g. weighting, them differently. In the paper at hand, we leave these tasks to future work.

4 Illustration Scenario

In this section, we provide a prototypical implementation to automatically complement goal concept hierarchies with means-ends relations, i.e. relating goals to actions which potentially contribute to their accomplishment. Complementing goal concept hierarchies by means-ends relations widens the range of operations which can be applied, e.g. generating action sequences to support planning procedures. The advent of social media platforms allows us to automate the process of extracting means-ends relations from the web.

4.1 Extracting Means-Ends Relations

We utilize Yahoo!Answers (Y!A), a social media platform, as resource to extract candidate actions which contribute to health-related goals. Y!A mediates the process of people posing questions and of people answering questions. For our purposes,

[4] According to Harris' distributional hypothesis [15]

these question/answer pairs are of particular value since a question often represents a person's goal and its answers potentially contain means for the goal's accomplishment. To illustrate our idea, we review a question/answer pair from Y!A:

Question:	"How can I lose weight easily?"
Answer:	"The best way to lose weight or maintain a healthy weight is through changing your eating habits and exercising regularly."

We might learn from this example that actions such as "change your eating habits" or "exercise regularly" might contribute to the goal "lose weight easily". We utilize Y!A's API to programmatically retrieve answers to submitted goal queries. Goals are translated into questions, e.g. by prefixing "how to", and are sent to Y!A. Answers are collected and prepared for (i) pre-processing, (ii) pattern-based action extraction and (iii) post-processing described in detail in the following.

Pre-Processing	We assume that clauses are natural boundaries for singular actions. During the pre-processing step, we thus seek to identify all clauses in the Y!A answers. We pursue a rule based approach by compiling a list of clause delimiters such as "if", "and" or "\n", as well as punctuations. Eventually, the clause's constituents are part-of-speech tagged[5].
Pattern-Based Action Extraction	We use extraction patterns to identify candidate actions. Extraction patterns are a combination of an indicator phrase and a verb phrase. By examining a small set of answers, we manually compile a list of heuristic indicator phrases including "you should", "you have to", "try", "start" or "you can". If an indicator phrase precedes a verb phrase (determined by part-of-speech tag information), we add the verb phrase to the list of candidate actions.
Post-Processing	Sanitization steps are applied: (i) candidate actions with less than three tokens are removed, (ii) duplicate entries are removed and (iii) too general actions and actions without proper object are removed by blacklisting; the manually compiled blacklist contains entries such as "do it", "make sure" or "know how".

4.2 Complementing Goal Concept Hierarchies

To automate the ranking of candidate actions, we decide on a "wisdom of crowds" approach which is based on (web) statistics. If a candidate action often co-occurs with a human goal, we assume that it contributes to the goal's accomplishment and thus assign a high rank. To access web statistics, we issue phrase searches to the web using the Yahoo!BOSS API[6]. We construct query strings for goal/action pairs where we apply manual processing to increase the likelihood for hits. This manual processing includes correcting misspellings or adding personal pronouns at correct positions, e.g. "clean out my body". As ranking metric, we use an adapted version of point-wise mutual information [16]:

$$Score(goal, action) = \frac{Hits(action + "\ to" + goal)}{Hits(action) * Hits(goal)}$$

where we can eliminate the term Hits(goal) from the denominator as it is common to all candidate actions for a particular goal. In this implementation, we select the top

[5] Stanford log-linear part-of-speech tagger: http://nlp.stanford.edu/software/tagger.shtml.
[6] http://developer.yahoo.com/search/boss/

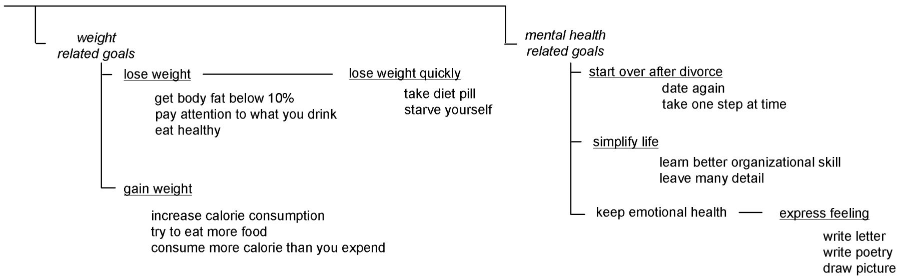

Fig. 3. Shows an excerpt of the golden standard, i.e. weight and mental health related goals. Underlined goal concepts are complemented with action concepts.

three ranked candidate actions[7]. Actions are conceptualized as it is illustrated with goal concepts in Section 2.1. Fig. 3 shows an excerpt of the golden standard which is complemented with automatically extracted action concepts.

The presented approach is capable of complementing goal concept hierarchies with reasonable action concepts (cf. Fig. 3). However, manual assessment of extracted and highly ranked action candidates reveals that the approach is vulnerable to false positives (due to the lack of human judgment). Moreover, as we use Y!A we assume that goals (as part of a question) have already been asked and answered. Hence, we were able to extract actions only for a third of all goals. To find more goal/action pairs, other social web sources such as how-to sites appear to be promising. Lastly, we take a closer look at goals without extracted actions. While some goals are simply not contained in Y!A such as "recover from open reduction ankle surgery" or "make sure skin heals after cryosurgery", we hypothesize that Question/Answer systems are not suitable for "find" goals, e.g. "find vascular surgeon", and "buy" goals, e.g. "buy weight watchers food online". Specialized services on the web, e.g. identified by semantic retrieval [7], appear to be more suitable means to accomplish these goals.

5 Related Work

In this section, we review work from two relevant research areas: (i) commonsense knowledge acquisition and (ii) goal concept hierarchies.

Commonsense Knowledge Acquisition: In [17], Minsky lays out a vision of machines that are capable of behaving intelligently. Realizing this vision requires the acquisition of real-world knowledge including, but not limited to, knowledge about human goals. Understanding human goals can, e.g. help to answer *why* questions about user behavior and user interactions (cf. [4]), to reason about people's goals or to generate action sequences that implement goals (planning, cf. [1], [18]). Common sense knowledge comprises fact-based knowledge as well as knowledge about other aspects including emotions, temporal contexts or human goals (cf. [19]). It is assumed that every person possesses commonsense knowledge which spans a broad spectrum

[7] Other strategies are conceivable as well, e.g. introducing a threshold.

of human experiences. Cyc [20] and ConceptNet [19] are ongoing research projects aiming to capture commonsense knowledge including knowledge about human goals. By studying approaches to automatically construct goal concept hierarchies, this work contributes to further explicate knowledge about human goals and thus making it accessible to and utilizable for machines.

Goal Concept Hierarchies: Concept hierarchies are mental constructs to organize knowledge. They are considered vital for knowledge-based systems since they allow for a concise and abstract representation of a domain. Hierarchies of goal concepts have proven valuable in several research areas including (i) web search, (ii) intelligent user interfaces or (iii) semantic task retrieval. (i) Every search query either explicitly or implicitly reflects a person's underlying search goal. By capturing these intentional structures in a hierarchy, a person's search goal can be better predicted and thus more relevant results can be retrieved. While Broder [2] presents a high-level categorization which is manually created, Yin et al. [3] seek to automate the construction process by using clustering algorithms. (ii) Intelligent user interface research benefit from better understanding relationships between people's goals and their actions explicated e.g. in a concept hierarchy. In [5], Liu's GOOSE system implements a goal-oriented search interface utilizes ConceptNet's [19] goal knowledge structures to reformulate queries. (iii) Growing demands of the mobile web community encouraged researchers to develop semantic retrieval mechanisms. The idea is to better support people's goals on the web by finding appropriate web services. To that end, Naganuma et al. [7] present a knowledge modeling framework which specifies the semantic description of task and goal knowledge. Based on this semantic description, we investigated the construction of hierarchies reflecting users' real world activities in [6]. In this work, we continue our research to automate the construction of goal concept hierarchies.

6 Conclusions

In this work, we study the process of automatically constructing goal concept hierarchies. Automating this process is relevant for knowledge engineers to address the knowledge acquisition bottleneck [12]. Using token-based features, the HAC algorithm performs best reaching taxonomic overlaps of up to ~51% for the entire domain and up to ~60% for individual health subdomains. Our study indicates that human participation is required in the construction process and that its extent has an effect on various characteristics including quality. Human participation comprises pre- and post-processing steps, e.g. to compile gazetteer lists or to devise extraction patterns.

For further automation, we suggest examining strategies from research areas such as (i) open information extraction (cf. [21]), e.g. to automatically devise patterns based on seed examples, or (ii) human computation (cf. [22]) and crowdsourcing approaches to utilize human resources for quality assessments.

Our findings are relevant for knowledge engineers interested (i) in acquiring knowledge about human goals as well as (ii) in automating the process of constructing goal concept hierarchies.

Acknowledgements. The Know-Center is funded within the Austrian COMET Program under the auspices of the Austrian Ministry of Transport, Innovation and Technology, the Austrian Ministry of Economics and Labor and by the State of Styria. COMET is managed by the Austrian Research Promotion Agency FFG.

References

[1] Carberry, S.: Techniques for Plan Recognition. Journal of User Modeling and User-Adapted Interaction 11(1-2) (2001)
[2] Broder, A.: A taxonomy of web search. SIGIR Forum 36(2) (2002)
[3] Yin, X., Shah, S.: Building taxonomy of web search intents for name entity queries. In: Proceedings of the 19th International Conference on World Wide Web (2010)
[4] Smith, D., Lieberman, H.: The Why UI: using goal networks to improve user interfaces. In: the 14th International Conference on Intelligent User Interfaces (2010)
[5] Liu, H., Lieberman, H., Selker, T.: GOOSE: A Goal-Oriented Search Engine with Commonsense. In: De Bra, P., Brusilovsky, P., Conejo, R. (eds.) AH 2002. LNCS, vol. 2347, pp. 253–263. Springer, Heidelberg (2002)
[6] Fukazawa, Y., Ota, J.: Automatic Modeling of User's Real World Activities from the Web for Semantic IR. In: Semantic Search Workshop (2010)
[7] Naganuma, T., Kurakake, S.: Task Knowledge Based Retrieval for Service Relevant to Mobile User's Activity. In: Gil, Y., Motta, E., Benjamins, V.R., Musen, M.A. (eds.) ISWC 2005. LNCS, vol. 3729, pp. 959–973. Springer, Heidelberg (2005)
[8] Pass, G., Chowdhury, A., Torgeson, C.: A picture of search. In: Proceedings of the 1st International Conference on Scalable Information Systems (2006)
[9] Strohmaier, M., Kröll, M.: Acquiring knowledge about human goals from search query logs. Information Processing and Management (2011)
[10] Cimiano, P., Hotho, A., Staab, S.: Learning concept hierarchies from text corpora using formal concept analysis. Journal of Artificial Intelligence Research 24, 1 (2005)
[11] Maedche, A., Staab, S.: Measuring similarity between ontologies. In: Proceedings of the European Conference on Knowledge Engineering and Management (2002)
[12] Lieberman, H., Smith, D., Teeters, A.: Common consensus: a web-based game for collecting commonsense goals. In: Proc. of the Workshop on Commonsense and IUI (2007)
[13] Havasi, C., Speer, R., Alonso, J.: ConceptNet 3: a Flexible, Multilingual Semantic Network for Common Sense Knowledge. In: Advances in Natural Language Processing (2007)
[14] Cimiano, P.: Ontology Learning and Population from Text (2006)
[15] Harris, Z.: Mathematical Structures of Language. Wiley, New York (1968)
[16] Turney, P.D.: Mining the Web for Synonyms: PMI-IR versus LSA on TOEFL. In: Flach, P.A., De Raedt, L. (eds.) ECML 2001. LNCS (LNAI), vol. 2167, pp. 491–502. Springer, Heidelberg (2001)
[17] Minsky, M.: Commonsense-based interfaces. CACM 43, 8 (2000)
[18] Tenorth, M., Nyga, D., Beetz, M.: Understanding and Executing Instructions for Everyday Manipulation Tasks from the WWW. In: Intern. Conf. on Robotics and Automation (2010)
[19] Liu, H., Singh, P.: ConceptNet - A practical commonsense reasoning tool-kit. BT Technology Journal 22(4) (2004)
[20] Lenat, D.: Cyc: a large-scale investment in knowledge infrastructure. CACM 38(11) (1995)
[21] Banko, M., Cafarella, M., Soderland, S., Broadhead, M., Etzioni, O.: Open information extraction from the web. In: Proceedings of the 20th International Joint Conference on Artifical Intelligence (2007)
[22] Van Ahn, L.: Human Computation. PhD Thesis. Carnegie Mellon University (2005)