Incremental Analogical Learning with Qualitative Representations of Quantity

Will Hancock[1], **Walker Demel**[2], **Kenneth D. Forbus**[2]
Northwestern University

wwhancock@u.northwestern.edu,{walker.demel,forbus}@northwestern.edu

Abstract

Quantities are ubiquitous in our conceptualization of the world, and the ability to learn and reason with them is an important aspect of commonsense reasoning. Existing cognitive models of similarity and generalization often lack sensitivity to quantitative knowledge, and those that are often represent it implicitly, meaning that it is not available for further learning or reasoning. This paper presents an extension to analogical reasoning processes that enables learning from mixed qualitative and quantitative knowledge. This is accomplished by utilizing qualitative representations of quantity, and by leveraging structure mapping to build schemas incrementally, maintaining probability distributions for quantitative knowledge, and then using these distributions to generate predicates that participate in structured generalization. This extension, called **A**nalogical**Q**uantity**E**stimation (AQE) is both incremental and unsupervised, and our results show that AQE performs significantly better than a baseline where quantitative knowledge is not taken into account. In addition, we compare AQE to a standard linear regression estimator, which, despite being batch and supervised, does not perform significantly better than AQE, and in some cases, performs worse.

Introduction

Commonsense knowledge is playing an increasingly important role in the development of AI systems. Many large-scale knowledge bases are emerging that encode general facts about the world using both structured qualitative and quantitative knowledge. Such knowledge is available in large open-domain knowledge bases such as *OpenCyc*, *DBpedia* and *WikiData*.

The ability to learn and generalize from these knowledge sources is therefore useful to any AI agent. Most existing computational models of retrieval and similarity cannot use numerical representations (Forbus et al., 2017; Holyoak and Thagard, 1989; Hummel and Holyoak, 1997), leading to quantitative information being ignored in computation of similarity. There are models in case-based reasoning (Ram and Santamaria, 1997) that use numeric information, but they employ ad hoc similarity metrics that are not psychologically grounded. A major motivation of this work is to generate cognitively plausible symbolic representations of quantity and show that these representations aid in learning.

In this paper, we introduce a novel algorithm, AQE, that improves an existing analogical learner so that it is sensitive to quantity. A similar idea was proposed by Paritosh (2004), which introduced a computational model called *CARVE*. AQE extends CARVE in two ways. First, CARVE's quantity symbolization was external to the analogical learner and needed to be run manually. Second, this symbolization process was batch, meaning that it needed access to an entire dataset before learning, and symbolization needed to be complete before any learning took place. AQE addresses these issues by automatically symbolizing quantities incrementally as new cases are generalized. Additionally, CARVE did not find any regularities in the data it was tested on, whereas our model shows significant improvement over a baseline.

AQE is tested by estimating quantities for two datasets derived from Wikidata; one containing knowledge about countries, and the other knowledge about universities. Wikidata contains vast amounts of knowledge in a wide array of domains, and therefore is a useful resource that contains a wealth of ground facts that can be used for commonsense reasoning (Forbus and Demel, 2022).

We begin by introducing the most relevant related work on systems that used mixed qualitative and quantitative knowledge. Then we present AQE, including the qualitative representation scheme and its incorporation into an existing analogical learner. Finally, we show results for experiments on two Wikidata datasets, ending with conclusions and future work.

Related Work

We give a brief overview of previous computational models that use mixed qualitative and quantitative representations, as well as related models of similarity and retrieval.

Computational Models

There are many examples of representational schemas that combine structured and quantitative knowledge. Hinton's (1979) model of mental imagery combines structured knowledge with numerical properties. Both ACT-R (Anderson, 2009) and SOAR (Laird, 2012) use numerical components in their representations, for example, statistical metadata on recency, frequency, and utility for symbolic structures. There are currently several theoretical frameworks that tightly integrate logic and probability, including Markov Logic Networks (Richardon and Domingos, 2006), while Rosenbloom's (2013) SIGMA cognitive architecture is exploring how to use graphical models to build a complete cognitive architecture, including both symbolic and statistical reasoning.

Many of these models treat quantity implicitly, meaning that it is not available at the level of knowledge. On the other hand, explicit reification is useful because it allows for graceful extension in learning and reasoning, as well as access to the richer semantics of quantity ontologies, such as QP theory (Forbus, 2019).

In addition, these models often require batch learning, which is problematic for cognitive agents because all previous knowledge must be stored. On the other hand, AQE incrementally accumulates distributional knowledge over quantities, meaning that distributions can be updated online as new examples are generalized.

There is converging psychological evidence for structured models of retrieval, similarity, and generalization. One limitation of existing models of analogical processing, e.g., ACME (Holyoak and Thagard, 1989), LISA (Hummel and Holyoak, 1997), ABSURDIST (Goldstone and Rogosky, 2002) is that they do not handle numerical properties adequately. In most of these models, numbers are treated like symbols, so 99 and 100 are as similar/different as 99 and 10000. AQE addresses this issue by automatically symbolizing quantity using the qualitative representations proposed by CARVE, creating new predicates that contribute to similarity in analogical learning.

Background

Next, we overview the analogical learning stack (SME, MAC/FAC, and SAGE) that we are extending, CARVE, a computational model of quantity estimation that we are building on top of, and Wikidata, the source of our data.

Analogical Learning

The generalization mechanism for AQE is built on models inspired by Gentner's structure-mapping theory of analogy and similarity (Gentner, 1983). AQE uses the Structure Mapping Engine (SME; Forbus et. al, 2016) for analogical matching, MAC/FAC for retrieval, and SAGE for analogical generalization. These analogical processes have been used in a wide range of domains, including sketch recognition (Chen et al., 2023), learning to play strategy games (Hancock and Forbus, 2021), and question answering (Crouse et al., 2019), and so we hypothesize that it will be useful for learning with representations of quantity. We summarize each component in turn.

The structure mapping engine (SME) is a domain-general computational model of analogy and similarity, based on Gentner's structure mapping theory. It returns a set of mappings between a base and a target, both structured representations, along with a similarity score for each mapping. Each mapping contains (1) correspondences that map entities and expressions in the base with entities and expressions in the target, (2) a numerical structural evaluation score of the quality of the mapping, and (3) candidate inferences. Candidate inferences are expressions that occur in the base description and not in the target but can be hypothesized to hold in the target.

The MAC/FAC algorithm (Forbus, Gentner, and Law, 1995) is a model of analogical retrieval. MAC/FAC takes as input a probe description (a set of facts) and a set of examples, and returns the example that is most similar to the probe. MAC/FAC stands for many are called, few are chosen. Retrieval is a two-stage procedure. In the MAC stage, each case is represented by a *content vector*. Each dimension in a content vector represents a predicate, and its magnitude corresponds to the number of occurrences of the predicate in that case. The dot product of two content vectors provides a rough estimate of what SME would compute for a similarity score for the corresponding structured representations. This dot product is used as a pre-filter to reduce the number of comparisons made in the FAC stage, which are computationally more expensive. The MAC stage is a map/reduce operation, where a dot product for a content vector of the probe is computed in parallel with the vectors for all items in the case library, with the top three scoring cases passed on to the FAC stage. The FAC stage also is map/reduce but using SME on the probe and the three retrieved cases, keeping the best. The MAC stage provides scalability, since vector dot products are quite fast. The FAC stage provides the sensitivity to structure that human retrieval demonstrates.

The Sequential Analogical Generalization Engine (SAGE; Kandaswamy & Forbus, 2012) is a model of analogical generalization. SAGE learns models of concepts, incrementally, from examples. In SAGE, *generalization pools*, or *gpools*, are used to build up models of concepts. The number of gpools used for learning is determined by the number of concepts in a domain and the learning goals that arise. A gpool is subdivided into clusters of similar examples, or *generalizations*, and outliers that are not similar to any other cases or generalizations. Each generalization can be thought of as a component of a disjunctive model for the concept. In this sense SAGE is like k-means with outliers, except that there is no a priori determination of the number of clusters; the algorithm derives that from the data.

Generalization with SAGE involves assimilating new examples into gpools, and inference involves finding a generalization (or outlier) that is most similar to a probe case. For assimilation, an incoming case is used to retrieve existing outliers and generalizations within a gpool, using MAC/FAC. If the case is sufficiently similar to an existing generalization or outlier, as determined by a fixed assimilation threshold, it is merged with that item and a mapping is returned. Otherwise, a new outlier is created. If merging occurs between the probe and an outlier, then a new generalization is created.

In the case where two items are merged, SAGE uses information computed in a mapping to store metadata about the generalization. Probabilities are updated for aligned facts, reflecting the frequency of that fact within the generalization. For example, facts about international organization membership are included in each country case. After a number of country cases have been assimilated, a generalization will have a lifted facts corresponding to these memberships, and a probability for each fact. For example,
((MemberOfInternationalOrgFn
AllianceofSmallIslandStates) *<?country>): 0.96*

reflects the fact that a country is a member of the Alliance of Small Island States, and has a probability of .96 within the context of one generalization, meaning that 96% of the constituents of that generalization exhibit this attribute.

The <?country> placeholder is a skolem (new unique symbol) that is denoted in knowledge by a non-atomic term (*GenEntFn*). Probabilities for generalizations are updated every time a new example is assimilated. Statements whose probabilities become too low are eventually deleted, based on a fixed probability cutoff threshold.

Quantity Representation in CARVE

AQE builds upon representations of quantity and a computational model, CARVE, developed by Paritosh (2004). CARVE used two distinctions for representation of quantity: *distributional* and *structural* partitions. Distributional partitions map a continuous value to some ordered interval within a probability distribution. More than just the norm, ordered partitions can be defined within the distribution (e.g. small, large) for many quantities, which are construed as a qualitative decomposition of the space. There is psychological evidence that suggests that we can and do accumulate distributions of quantities (Malmi and Samson, 1983; Fried and Holyoak, 1984; Kraus et al, 1993). Distributional partitions are represented by statements of the form

(isa *<?country>* (<?amount> *<?qtype>*))

For example, the USA has a high literacy rate relative to all other countries in the world, represented by:

(isa USA (HighAmountFn LiteracyRate))

Whereas distributional partitions decompose individual quantities, structural partitions highlight how quantities are constrained by what values other quantities in the system take. For instance, GDP tends to increase as a country's population increases, and literacy rates tend to increase with GDP. These constraints represent the underlying mechanisms, or correlations within the domain. *Limit points* decompose values into regions where the underlying correlational story is different (e.g., rich vs poor nations), which induces important and interesting distinctions of quality on the space of quantity.

Wikidata

The AQE algorithm is domain-independent and ontology independent. This work focuses on readily available structured knowledge derived from the Wikidata dataset. Wikidata is a collaborative knowledge graph that serves as a repository of structured data for a wide range of information from many different domains. Like its sibling Wikipedia, Wikidata utilizes the distributed-community model of editors—as of this writing, thousands of editors and bots have made over 1.6 billion edits to over 97 million items. This model allows Wikidata to serve as the downstream aggregate of otherwise independent structured data sources.

Wikidata is organized around items, with each having a unique identifier (QID) and a set of statements about it. Each statement is an RDF triple of <subject property value>. For example, "the United States is a member of the World Health Organization" can be expressed as *<member of (P463), United States (Q30),World Health Organization (Q7817)>* where the terms in italics are the English rendition of the objects whose ids are in parentheses. In QR terminology, items are entities and values are quantity values. In some cases the quantity type is obvious (e.g. Area, Color), while in others it is more opaque (Statistical Population). Any property can specify constraints on its value. Certain properties specify that their values must be a string, number, date, URL, media file, or another Wikidata entity. Other properties, like capital (P36) enforce no more than one value since most states have only one capital. Since Wikidata consists of RDF triples, it can be queried via a SPARQL endpoint (query.wikidata.org). In Wikidata, predicates like point in time (P585) can be used to qualify statements like population (P1082), for which there may be several different assertions that hold in different years. In the case where a country's capital (P36) may have changed, values can be associated with a start time (P580) and end time (P582).

Analogical Quantity Estimation

Recall that SAGE computes progressive structural overlap over incoming cases, resulting in a set of disjunctive generalizations for a concept. For example, in this work a generalization might denote the set of wealthy European nations. In this sense, generalizations can be viewed as structural partitions that describe some latent concept (i.e. rich countries). The goal of structural partitioning is to assign cases to generalizations that correspond to useful distinctions (for instance, groups of developed and underdeveloped nations). Learning for AQE consists of two steps. In the first, quantitative facts are symbolized; that is, continuous quantities are mapped to qualitative distributional partitions, and the resulting new facts are added to the original case. In the second step, this augmented case is added to a separate gpool, which learns structural partitions in the data. We outline each of these steps next.

Distributional Partitioning

The first step for AQE is to encode numeric facts in incoming cases. Many quantity estimators, e.g. regression, assume that

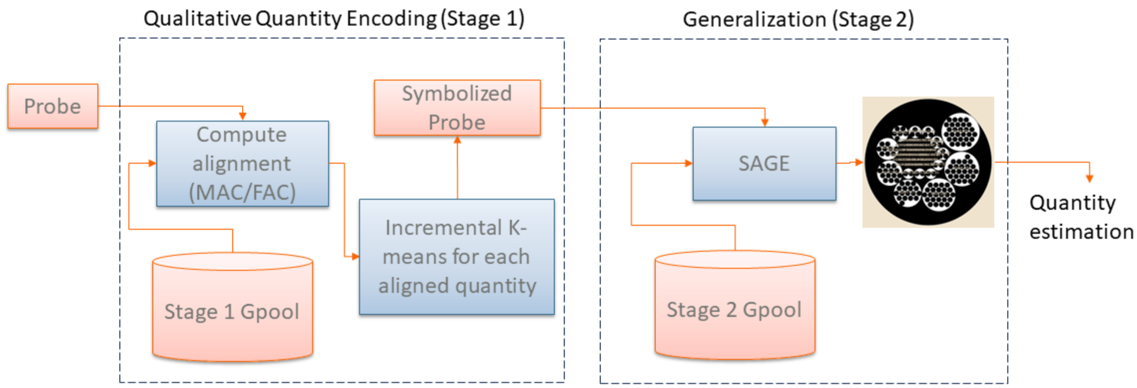

Figure 1: Overview of the AQE encoding, generalization, and inference processes. Stage 1 encodes qualitative facts for each numeric quantity found in a case. This is achieved by accumulating statistical information about each quantity type in the stage 1 gpool. For example, country GDP will have an associated k means. This distributional knowledge is used to encode quantitative knowledge in incoming cases. Stage 2 generalizes the newly symbolized cases, resulting in a set of generalizations (structural partitions), each accumulating statistical information about constituent cases. This model is then used for inference, to estimate quantities for new cases.

incoming data is unstructured, and that attributes are already aligned. Since Wikidata combines structured and unstructured knowledge, this poses an additional challenge to learning. That is, entities and attributes must be aligned before learning can take place. For learners like regression, this is handled outside of the learning mechanism, often manually. One benefit of AQE is that this procedure is handled automatically by computing analogical mappings and is tightly integrated into the learning mechanism. Thus, the first step in symbolizing quantities is to compute attribute alignments (Figure 1, qualitative quantity encoding). Once this is complete, distributions for aligned quantities are used to map continuous quantities to distributional partitions, which we describe next.

First, an incoming quantity must be mapped to a set of previously seen quantities. For example, to symbolize the literacy rate of the USA, which is 99.4 as of 2022, then this quantity should be compared with the distribution for literacy rates of all previously seen countries. This is handled with SAGE by maintaining a gpool that has an assimilation threshold of zero. Recall that the assimilation threshold sets the minimum requirement for two cases to be considered similar. An assimilation threshold of zero means that a gpool will have a single generalization that contains all assimilated cases. While not useful for learning (because it makes no distinctions), this model is useful because it provides a global schema. This schema provides useful metadata about facts in the dataset. For one, the relative frequency of each aligned fact is stored, (e.g. 3.4% of countries border Cameroon). Second, it associates each quantity type with information about the values that that quantity type has taken. The goal is to separate each quantity type into a predetermined number of qualitative partitions. This is achieved with an online k-means algorithm. Given an unseen quantity, it is assigned to one of the K distributions by minimizing the Euclidian distance between the quantity and the norms of each distribution. If less than k quantities have been seen, a new distribution is created, and the new quantity is set as the mean.

For this paper, the number of distributional partitions is set at five, as we have found that this is a good balance between expressiveness and relevance. Too expressive (too many partitions) and all quantities tend towards dissimilarity. Too few distinctions, and all quantities tend towards similarity. Using five partitions results in a quantity space that can be interpreted as (very small, small, medium, large, very large). The number of distributions K can be set at the level of a gpool by by asserting a fact

(kMeansForQuantityAnalysis <?gpool> <?K>)

in the knowledge base.

The next step is to generate qualitative facts based on the assignment of a quantity to one of the K distributional partitions. If fewer than K quantities have been seen, then no fact is generated. Otherwise, a new fact is created, e.g.

(isa USA (HighAmountFn GDP))

and added to the existing case in place of the prior quantitative fact.

Next, the associated distribution is updated to reflect the new quantity. SAGE stores with each distribution a set of statistics: the cardinality, minimum, maximum, mean, and sum of squared error of the constituent quantities. This metadata is used later on for inference, which is detailed below.

Structural Partitioning

Once quantities in a case have been symbolized, the case is given as input to a second SAGE gpool, this time with a nonzero assimilation threshold. The assimilation threshold used for these experiments is .6, which is a standard value that has been successful for learning across many domains. In this step, both existing qualitative as well as the new symbolized

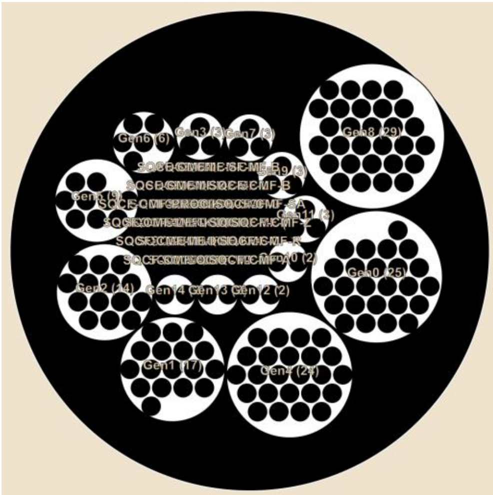

Figure 2: A SAGE Gpool consisting of 15 generalizations (white circles), each containing individual cases (black dots). Generalizations represent groups of similar cases (structural partitions of the dataset).

qualitative facts are taken into account by analogical matching. Figure 2 shows an example gpool with white circles designating generalizations containing similar countries. Each generalization reflects some structural partition in the source dataset. Structural partitions are a reflection the system's understanding of the correlational structure of a set of cases.

The gpool for the second stage accumulates the same statistical knowledge about quantity distributions as the first stage, over quantities of cases within the same generalization. In the next section, we describe how, along with analogy, these statistics contribute to inference in AQE by allowing quantity estimation for quantities in held out cases.

Quantity Estimation

For inference, the idea is to estimate an unseen quantity for some case. First, knowledge about the target quantity is removed from the case. Then, quantity estimation proceeds by first symbolizing all quantities, using the gpool from stage 1. First SAGE retrieves a mapping between the probe case and a generalization. Since the stage 1 gpool has an assimilation threshold of zero, all cases are similar to the single generalization, and a mapping is guaranteed. This mapping aligns quantities in the probe to previously seen quantities from training. For each aligned quantity, the k-means algorithm assigns it to one of K distributions. This assignment is used to generate a qualitative fact, as detailed previously. This fact is added to the probe case, and once this has been performed for all aligned quantities, inference proceeds to stage 2.

Next, SAGE retrieves a mapping from the augmented case to an object in the stage 2 (structural) gpool. If a mapping to a generalization is found, the mean of the target quantity type for that generalization is used as the estimate. If no mapping is found, then the estimate is the marginal average for that quantity type across all cases in the gpool. If the case maps to an outlier, then the quantity from the outlier is used for prediction, or the marginal over all cases in the gpool if the outlier does not have a quantity value for that quantity type.

Evaluation

This paper evaluates AQE on a set of cases that were extracted from Wikidata. Next, we describe this extraction procedure and then discuss how the resulting dataset is used to evaluate AQE.

Case Construction from Wikidata

For learning in AQE, we translate from Wikidata to the open-license knowledge base NextKB[1], which is used for AQE experiments. Data from Wikidata was pulled using the public SPARQL endpoint at query.wikidata.org. For the country dataset, ten quantitative attributes were queried for the year 2022 (population, GDP, GDP per capita, median income, democracy index, life expectancy, fertility rate, area, literacy rate, and human development index) and 6 qualitative attributes (continent, bordering countries, bordering bodies of water, language(s) spoken, international organization memberships, and currency). Overall, 197 cases were generated, having an average of 40 facts each.

For the set of university cases, qualitative attributes are (instanceOf; P31), organizational memberships (P463), and Carnegie Membership Classification (P2643). For quantities, students count (P2196), total assets (P2403), employees (P1128), admission rate (P5822), endowment (P6589), and admission yield rate (P10263) were used. For each university, the latest available quantity for each quantity type was used. All quantities are from 2019 and later, up to the year 2023. This dataset was extracted on July 31, 2023. Cases were generated for universities that were founded prior to 1860, which resulted in 231 university examples. Those that did not have any associated quantitative knowledge were removed, resulting in 194 cases.

Attributes for cases were chosen based on the hypothesis that there is a rich underlying correlational structure that can be learned. These facts were translated into OpenCyc's ontology language for use within NextKB. For some predicates there was a natural correspondence, such as nominal GDP in Wikidata and grossDomesticProduct in OpenCyc. Other predicates were missing from OpenCyc and thus hand ontologized, e.g. human development index as the predicate hdiOfCountry and percentage of applicants admitted as percentApplicantsAdmitted.

For example,

<United States (Q30), population (P1082), 331,449,281>

[1] https://www.qrg.northwestern.edu/nextkb/index.html contains downloadable files in various formats, browsers, and reasoning systems. It uses Creative Commons Attribution 4.0 licensing, compatible with OpenCyc, FrameNet, and other resources.

results in

(populationOfRegion
UnitedStatesOfAmerica
((UnitOfCountFn Person) 331449281)).

Experiment

For the experiments, AQE is evaluated against two baselines: (1): analogical quantity estimation without quantity symbolization and (2): against a standard linear regression estimator. Our hypothesis is that AQE will outperform analogical quantity estimation without qualitative representations of quantity. Additionally, results from a standard linear regression estimator are included. Recall that AQE is both incremental and unsupervised; incremental learners are well known to lack statistical guarantees of their batch counterparts due in part to the stochastic effects of initialization. The results for linear regression are included as a means of comparing learning performance of AQE vs a technique with better learning guarantees.

To run the experiments, standard cross validation is used to partition each dataset into ten folds, each consisting of a train and test set. For countries, there are 197 cases, and 194 for universities, resulting in a test set for each of the ten folds consisting of approximately 20 cases for both datasets. The folds are generated by first randomizing the cases, and then generating ten partitions based on the ordering from this randomization. AQE and the incremental baseline are implemented in Allegro Common Lisp 10.1. A seed for the random state in Allegro Common Lisp is set to 55 for the baseline and AQE conditions, as well as generation of the cross validation set. The linear regression estimator is run using the implementation in Python's ScikitLearn module, using default parameters, also using the same cross validation set that was generated in Allegro Common Lisp. Learning regression models requires vectorizing structured knowledge from each dataset. This is accomplished by manually creating a mapping, where each quantity type is considered a feature, and each unique qualitative attribute (e.g. currency, organization membership) is represented by a one-hot vector. Missing quantities are imputed using Python's impute function from the SciPy module. This results in 883 features across the 197 country cases for universities, and 181 features across 194 cases for the university dataset.

For the country dataset, each condition is tested on four different quantity types: life expectancy (**LE**), human development index (**HDI**), democracy index (**DI**), and nominal GDP (**GDPnom**). For universities, average yield percent (**AYP**), percent applicants admitted (**PAA**), and number of emplyees (**NOE**) are tested.

Results

For countries, our results show that AQE performs significantly better ($p < .05$) than the baseline for every quantity that was tested. Additionally, the regression condition fails to perform significantly better than AQE for any quantity ($p > .05$), and AQE outperforms the regression estimator in one instance (nominal GDP).

	LE	**HDI**	**DI**	**GDPnom**
baseline	60.84	.023	5.15	5.54e8
AQE	20.8	.0088	2.63	**3.79e8**
regression	**19.88**	**.0062**	**1.95**	4.9e8

Table 1: mean squared error across 10 folds for four quantity types (life expectancy, human development index, democracy index, and nominal GDP) across three experimental conditions.

For the university dataset, all experiments were run using the same parameters that were used for the country cases. We tested AQE on average yield percent **(AYP)** (the percentage of students that enroll given acceptance), percent applicants accepted (**PAA**), number of employees (**NOE**), and endowment value (**EV**). For this experiment, AQE outperformed the incremental baseline as well as regression for all quantities tested. Admission yield percent and percent applicants admitted showed significant improvement over the incremental baseline ($p < .05$). The regression condition suffered due to overfitting on certain folds, resulting in large out-of-distribution predictions.

	AYP	**PAA**	**NOE**	**EV**
baseline	.017	.059	1.31e7	3.9e19
AQE	**.011**	**.034**	**8.46e6**	**2e19**
regression	39443	7.5	9.43e9	2.95e19

Table 2: mean squared error across 10 folds for four quantity types (admission yield percent, percent applicants admitted, number of employees, and endowment value).

Explainability

The qualitative representations of quantity used in AQE also result in explainable models, because they are compatible with natural language. The final learned model (stage 2) represents a disjunction over structural partitions of the data. Figure 3 shows a subset of facts from one of these learned structural partitions. In SAGE terms, this corresponds to a

`Fact`	Prob
((MemberOfInternationalOrgFn AfricanDevelopmentBank) *<?country>)*	1.0
((MemberOfInternationalOrgFn AfricanUnion) *<?country>)*	1.0
((MemberOfInternationalOrgFn InternationalBankforReconstructionandDevelopment) *<?country>)*	1.0
(isa *<?country>* (CountryTypeFn (VeryLowAmountFn grossDomesticProduct-Nominal)))	.862

Figure 3: Example facts from a single generalization (structural partition) with 29 member cases

generalization, which stores probabilities of individual facts. The depicted generalization in Figure 3 shows that every constituent is a member of the African Development Bank. Furthermore, inspection of the model reveals that 86.2% of the participants have a very low GDP. The nature of these representations means that they are inspectable. In the next section, we discuss possible extensions that use these learned models for further learning.

Discussion and Future Work

This paper introduces AQE, an extension to SAGE that enables learning with quantitative knowledge by automatically symbolizing those quantities into predicate statements that denote distributional partitions. Furthermore, the experiment shows that these representations can assist in estimating quantity by using analogy to learn salient structural partitions of the underlying data in two datasets. Specifically, using qualitative representations significantly improves over a baseline in which these representations are not included. AQE is also compared against a linear regression estimator, which, despite being supervised and batch, does not perform significantly better than AQE in the first experiment, and in the second experiment, performs worse for every quantity.

As Paritosh (2004) points out, relative magnitudes such as *large* are context dependent and thus elude global definition. A person might be tall with respect to the general populace, but short compared to the set of professional basketball players. These ecological constraints surrounding judgements of this kind mean that the ability to quickly estimate from a few examples is especially useful, because new contexts are encountered frequently. The relative data efficiency of AQE is a boon compared to the data-hungry nature of many current statistical learners. This raises the possibility of applying AQE to new domains, such as concept acquisition in situated learning (e.g. learning *near* and *far* by symbolizing quantity).

Furthermore, declarative representations like the ones used in this paper allow for extensibility, in that they can be used as a foundation for other kinds of reasoning. For example, further refinement is possible by connecting the learned representations with the semantics of a more expressive qualitative reasoning framework, e.g. Qualitative Process theory (Forbus, 2019). This opens up the possibility of refining these models by incorporating knowledge from other sources, e.g. language. One way that this might be accomplished is by leveraging causal relationships parsed from language to highlight what is salient for a given estimation task. The idea is that this causal knowledge could improve model accuracy by filtering out noise introduced by non-salient attributes, e.g., extending a quantitative anchoring framework such as KNACK (Paritosh and Klenk, 2006).

Acknowledgements

This research was sponsored by the US Air Force Office of Scientific Research under award number FA95550-20-1-0091.

References

Anderson, J. R. (2009). How Can the Human Mind Occur in the Physical Universe? Oxford University Press.

Chen, K., Forbus, K., Srinivasan, B. V., & Chhaya, N. (2023). Sketch Recognition via Part-based Hierarchical Analogical Learning.

Crouse, M., McFate, C., & Forbus, K. (2018). Learning From Unannotated QA Pairs to Analogically Disambiguate and Answer Questions. Proceedings of the AAAI Conference on Artificial Intelligence, 32(1).

Demel, W., Hancock, W., & Forbus, K. (under review). KNACK v2: Using Analogical Generalization over Qualitative Representations for Quantitative Estimation. *Under review.*

Forbus, K. D. (2019). Qualitative Representations: How People Reason and Learn about the Continuous World. MIT Press.

Forbus, K. D., & Demel, W. (2022). Integrating QR Quantity Representations with the Semantic Web: A Progress Report.

Forbus, K. D., Ferguson, R. W., Lovett, A., & Gentner, D. (2017). Extending SME to Handle Large-Scale Cognitive Modeling. Cognitive Science, 41(5), 1152–1201.

Forbus, K. D., Gentner, D., & Law, K. (1995). MAC/FAC: A model of similarity-based retrieval. Cognitive Science, 19(2), 141–205.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. Cognitive Science, 7(2), 155–170.

Goldstone, R. L., & Rogosky, B. J. (2002). Using relations within conceptual systems to translate across conceptual systems. Cognition, 84(3), 295–320.

Hancock, W., & Forbus, K. D. (2021). Qualitative Spatiotemporal Representations of Episodic Memory for Strategic Reasoning.

Hinton, G. (1979). Some Demonstrations of the Effects of Structural Descriptions in Mental Imagery.

Holyoak, K. J., & Thagard, P. (1989). Analogical Mapping by Constraint Satisfaction. Cognitive Science, 13(3), 295–355.

Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. Psychological Review, 104(3), 427–466.

Kandaswamy, S. and Forbus, K. (2012). Modeling Learning of Relational Abstractions via Structural Alignment. *Proceedings of the 34th Annual Conference of the Cognitive Science Society (CogSci).* Sapporo, Japan.

Kraus, S., Ryan, C. S., Judd, C. M., Hastie, R., & Park, B. (1993). Use of mental frequency distributions to represent variability among members of social categories. Social Cognition, 11(1), 22–43.

Laird, J. E. (2012). The Soar Cognitive Architecture. MIT Press.

Malmi, R. A., & Samson, D. J. (1983). Intuitive averaging of categorized numerical stimuli. Journal of Verbal Learning and Verbal Behavior, 22(5), 547–559.

Paritosh, P. K. (2004). Symbolizing Quantity. Proceedings of the Annual Meeting of the Cognitive Science Society, 26(26).

Paritosh, P. K., & Klenk, M. E. (2006). Cognitive Processes in Quantitative Estimation: Analogical Anchors and Causal Adjustment. The Proceedings of the 28th Annual Conference of the Cognitive Science Society, Vancouver.

Ram, A., & Santamaría, J. C. (1997). Continuous Case-Based Reasoning. Artificial Intelligence, 90(1–2), 25–77.

Richardson, M., & Domingos, P. (2006). Markov logic networks. Machine Learning, 62, 107–136.

Rosenbloom, P. (2013). The Sigma cognitive architecture and system.