Ontology for Water Spectroscopy Information Resources

A. Privesetsev , A. Fazliev, D. Tsarkov, J.Tennyson

Abstract—The ontology of the problems solution properties related to the water spectroscopy is described. These solutions were published in more than 600 articles during the last 50 years and are related to six typical molecular spectroscopy problems. The properties and their values characterizing solutions validity, root-mean-square deviations were represented as individuals of OWL-ontology. The results of selection rules verification in primary data sources are presented.

***Index Terms*—Scientific Annotation Model, Water Spectroscopy Ontology**

I. INTRODUCTION

In the last decade molecular spectroscopists have united to address key tasks requiring a cooperative solution. One of such tasks is the development of a comprehensive representation of the spectrum of the water molecule. A suitable theoretical strategy for representing the spectrum was formulated in the framework of two projects [2-4]. Firstly a protocol "Marvel", which is generally applicable in molecular spectroscopy, was formulated [4]. Secondly a collective effort was made at collecting and validating both calculated and measured spectral data [2]. Thus, the creation of prototypes of information systems on molecular spectroscopy accessible via the Internet was initiated.

At present we are completing the implementation of an information system (IS) on molecular spectroscopy [5, 6] with three-layer architecture [7]. In this architecture the knowledge layer contains ontology of the IS domain resources.

The usage of the knowledge layer is oriented on the classification of the data sources published in water spectroscopy. The classification is based on the groups of properties which describe the validity of the data sources and their pair correlations. Computer classification requires the logical inference over the ontology [8-10].

The results of the classifications are presented in the report.

II. MOLECULAR SPECTROSCOPY MODEL

The procedure of an ICS creation requires the formation of a domain model. In the most general case the domain description contains the concepts related both to its declarative and procedure knowledge parts. In this work we considered conceptualizations of two domains.

The first domain (molecular spectroscopy) contains the facts about the states and transitions in molecular gases. The concepts related to the description of physical system states are described in details while the concepts that characterize the processes of transition from one state into another are not described in details. *De facto*, such domain model is widely used, for example, in a series of physics domains in which the procedures of experimental results acquisition are well-established but the huge amount of data requires many years of experimental measurements.

The simplification of procedure knowledge is caused by the assumption due to the following fact: in practice, for example, in information description of molecular spectroscopy the most needed information is the quantitative information on molecules states. In molecular spectroscopy the procedure knowledge may be of interest only to a narrow circle of users that implement different methods of domain problems' solution.

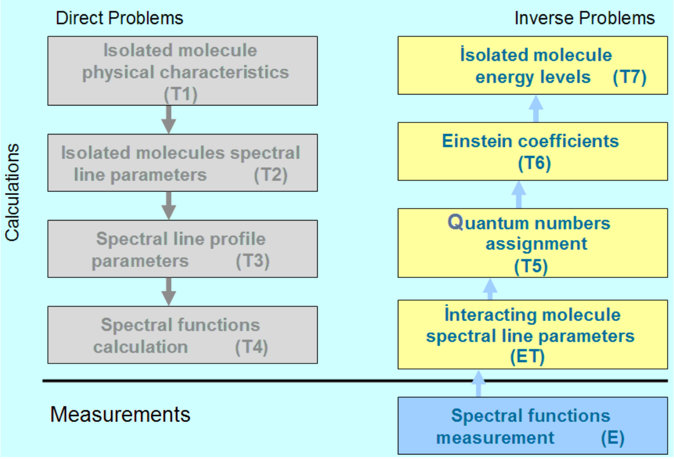
Direct Problems
Inverse Problems
Isolated molecule physical characteristics (T1)
Isolated molecules spectral line parameters (T2)
Spectral line profile parameters (T3)
Spectral functions calculation (T4)
Calculations
İsolated molecule energy levels (T7)
Einstein coefficients (T6)
Quantum numbers assignment (T5)
İnteracting molecule spectral line parameters (ET)
Measurements
Spectral functions measurement (E)

Fig. 1. Molecular spectroscopy model.

Another formal restriction caused by the necessity to represent knowledge on a computer is the choice of time-dependent entities model (occurents [11]) in which they are regarded as the formal problems described by IPO (Input-Processing-Output) – model. The degree of input and output data specification as well as the specification of tasks solutions methods for procedure knowledge model creation may be different. In this work the key component of an IPO task model is the output data (problem solutions). Input data and problems solution method in the framework of a given model relate to the properties of solutions of problem included in the scientific annotation of the resource.

Schematic model of molecular spectroscopy for the information sources of which the scientific annotations are composed automatically is presented in Fig. 1.

The second most important model is the model of information objects. In our work the *"annotation"* concept plays an important role in the conceptualization of this model. This concept describing the domain problem solution and its properties is described below.

This work was supported in part by the RFBR under Grant No. 08-07-00318-a.

A. Privesetsev and A. Fazliev are with the V.E. Zuev Institute of Atmospheric Optics, Tomsk, 634021 Russia (phone: 7-3822-492277, e-mail: faz@iao.ru).

D. Tsarkov is with Manchester University, Oxford Road, Manchester M13 9PL. (phone:+44-161-275-6817 e-mail: tsarkov@cs.man.ac.uk).

J. Tennyson is with the University College London, Gover St., London WC1E 6BU, UK (e-mail: j.tennyson@ucl.ac.uk).

III. SCIENTIFIC ANNOTATION MODEL

The data gathered in an information system which are presented in molecular spectroscopy as the solutions of the above problems should be described by annotations. In general the formation of an annotation depends on the problems and the methods of their solution in an information system. Problem digital solution is an information resource which has a typical part of annotation related to the properties of an abstract resource [1]. The Dublin Core properties may grant us the answers to the questions of authorship, publication place, rights and so on. The scientific part of the annotations is bound with formalized mathematical properties of the solution included in the resource.

In our opinion a scientific annotation should necessarily include information resource validity description. However there is one problem, i.e. the notion of validity has various interpretations in different domains. In molecular spectroscopy interpretation choice helps one to select an information source in the necessary frequency range or according to different validity criteria.

Finally an annotation may include the description of informational sources relations. In quantitative sciences a simple example described by such an annotation is the root-mean-square deviations of different solutions of one and the same problems.

For the purpose of classification scientific annotations are transformed into the individuals of applied ontology of spectroscopy problems. Such transformation is achieved with OWL DL ontologies specification language.

Molecular spectroscopy science and information resources may be divided into primary and composite. A primary resource in the chosen representation is a published problem solution of a domain related to the determined molecule.

A. Autonomous Part of Annotation

The given annotation part includes only the properties related to one solution of molecular spectroscopy task. It is presented in Fig.2 in the form of a structure including such individuals as “BroadeningSubstances”, “OutputData”, “Halfwidth_for_H2O”, “Shift_for_H2O” etc... All the individuals in Fig. 2 are marked as rectangles. Individuals are in bold, object properties are in common font and the properties of data type are in italic.

In the created annotations datatype properties are related to such properties as *hasReference* (string), *hasUncertainty* (boolean), *hasMaxWavenumber* (float) etc… The values of these properties are calculated automatically at the acquisition of data by the information system.

The data type properties are related to a series of validity criteria based on the checkup of the restrictions of physical quantities permitted values [8] in the domain. The condition necessary for the selection of such checkups in an automated information system is their computability and decidability. In most cases such conditions are met.

The validity criteria related to the domain of values of object properties and used in this work are related to existence restriction check. In the broad sense the existence of an information resource implies that it has URI [13] while in the narrow sense a scientific resource should be given only the URI from the list adopted by the scientific community. In this general meaning of physical quantity value existence determination is changed for a simpler task of checking the existence of parts of composite information resource among the earlier published problem solutions. Such criterion may be applied to the check of validity of composite problem solutions widely used in molecular spectroscopy (see, for example, [14,15]). Using this problem definition the proving of the completeness of a set of primary sources used in such a check becomes important. The formation of such a check was implemented in the framework of IUPAC [6] project including about 800 primary sources of water spectroscopy tasks solution (see Table 1).

Non-formalized restrictions defining the water spectroscopy tasks’ solutions validity criteria are used for the expert analysis of tasks’ solutions. Such analysis published in press is included into annotation. An example is the introduction of experts’ comments on incorrect attribution of quantum numbers in a series of works on the determination of transitions wave numbers for $H_2^{17}O$ and $H_2^{18}O$ molecules [2] to the annotations. Unlike bibliographic annotation the scientific annotation is the resource the contents of which changes with time and the majority of scientific annotation components are computable.

As the chosen annotation model allows us to automatically present an annotation in a form of an individual of a calculable and solvable applied ontology all the logical inferences following from the change of annotations’ properties are computable.

B. Nonautonomous Part of Annotation

In our opinion a scientific annotation may characterize not only a resource directly related to it but its relations to other resources as well. A common example typical for the description of water spectroscopy problems’ solutions is the usage of computable root-mean-square deviation of this solution with all the other solutions into the annotation. This part of an annotation allows us to make major conclusions about statistical correctness of the given solution.

In our model of this part of annotation maximal deviations on physical quantities’ values at the corresponding quantum numbers, the number of bands and the number of transitions or energy levels corresponding in a band are used along with integral value of root-mean-square deviation and the deviation of each of the corresponding bands.

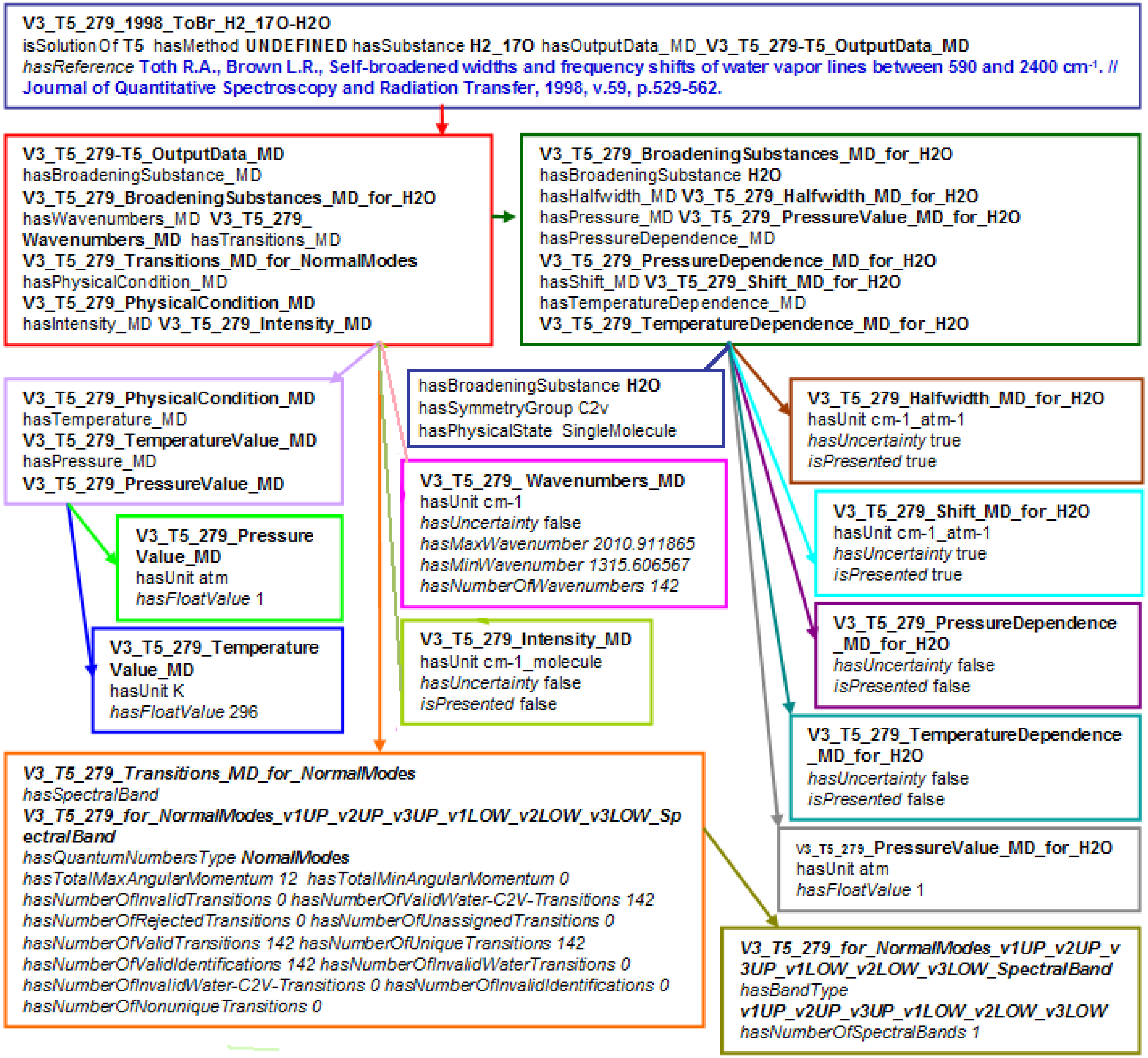
V3_T5_279_1998_ToBr_H2_17O-H2O
isSolutionOf T5 hasMethod UNDEFINED hasSubstance H2_17O hasOutputData_MD V3_T5_279-T5_OutputData_MD
hasReference Toth R.A., Brown L.R., Self-broadened widths and frequency shifts of water vapor lines between 590 and 2400 cm-1. // Journal of Quantitative Spectroscopy and Radiation Transfer, 1998, v.59, p.529-562.
V3_T5_279-T5_OutputData_MD
hasBroadeningSubstance_MD
V3_T5_279_BroadeningSubstances_MD_for_H2O
hasWavenumbers_MD V3_T5_279_
Wavenumbers_MD hasTransitions_MD
V3_T5_279_Transitions_MD_for_NormalModes
hasPhysicalCondition_MD
V3_T5_279_PhysicalCondition_MD
hasIntensity_MD V3_T5_279_Intensity_MD
V3_T5_279_BroadeningSubstances_MD_for_H2O
hasBroadeningSubstance H2O
hasHalfwidth_MD V3_T5_279_Halfwidth_MD_for_H2O
hasPressure_MD V3_T5_279_PressureValue_MD_for_H2O
hasPressureDependence_MD
V3_T5_279_PressureDependence_MD_for_H2O
hasShift_MD V3_T5_279_Shift_MD_for_H2O
hasTemperatureDependence_MD
V3_T5_279_TemperatureDependence_MD_for_H2O
V3_T5_279_PhysicalCondition_MD
hasTemperature_MD
V3_T5_279_TemperatureValue_MD
hasPressure_MD
V3_T5_279_PressureValue_MD
hasBroadeningSubstance H2O
hasSymmetryGroup C2v
hasPhysicalState SingleMolecule
V3_T5_279_Halfwidth_MD_for_H2O
hasUnit cm-1_atm-1
hasUncertainty true
isPresented true
V3_T5_279_Pressure
Value_MD
hasUnit atm
hasFloatValue 1
V3_T5_279_ Wavenumbers_MD
hasUnit cm-1
hasUncertainty false
hasMaxWavenumber 2010.911865
hasMinWavenumber 1315.606567
hasNumberOfWavenumbers 142
V3_T5_279_Shift_MD_for_H2O
hasUnit cm-1_atm-1
hasUncertainty true
isPresented true
V3_T5_279_PressureDependence
_MD_for_H2O
hasUncertainty false
isPresented false
V3_T5_279_Temperature
Value_MD
hasUnit K
hasFloatValue 296
V3_T5_279_Intensity_MD
hasUnit cm-1_molecule
hasUncertainty false
isPresented false
V3_T5_279_TemperatureDependence
_MD_for_H2O
hasUncertainty false
isPresented false
V3_T5_279_Transitions_MD_for_NormalModes
hasSpectralBand
V3_T5_279_for_NormalModes_v1UP_v2UP_v3UP_v1LOW_v2LOW_v3LOW_SpectralBand
hasQuantumNumbersType NomalModes
hasTotalMaxAngularMomentum 12 hasTotalMinAngularMomentum 0
hasNumberOfInvalidTransitions 0 hasNumberOfValidWater-C2V-Transitions 142
hasNumberOfRejectedTransitions 0 hasNumberOfUnassignedTransitions 0
hasNumberOfValidTransitions 142 hasNumberOfUniqueTransitions 142
hasNumberOfValidIdentifications 142 hasNumberOfInvalidWaterTransitions 0
hasNumberOfInvalidWater-C2V-Transitions 0 hasNumberOfInvalidIdentifications 0
hasNumberOfNonuniqueTransitions 0
V3_T5_279_PressureValue_MD_for_H2O
hasUnit atm
hasFloatValue 1
V3_T5_279_for_NormalModes_v1UP_v2UP_v3UP_v1LOW_v2LOW_v3LOW_SpectralBand
hasBandType
v1UP_v2UP_v3UP_v1LOW_v2LOW_v3LOW
hasNumberOfSpectralBands 1

Fig. 2. Computer processed structure of the individual that characterizes the properties of the solution of an inverse problem on spectral line profile parameters.

If a set of solutions includes all the known solutions of an inverse task then the case of zero value of root-mean-square deviations of the given source with all the other ones correspond to a case when the given source contains wave number values which were not measured or identified earlier.

Fig. 3 illustrates the structure of an applied ontology individual which is the representation of an information resource annotation on the logical level. The number of statements that characterize the relation of the solutions depends on the number of corresponding bands. Describing a pair correlation the total number of statements is defined by the following formula:

$$S = 131 + 5N_{bands},$$

where N_{bands} is the number of corresponding bands at the calculation of root-mean-square deviation.

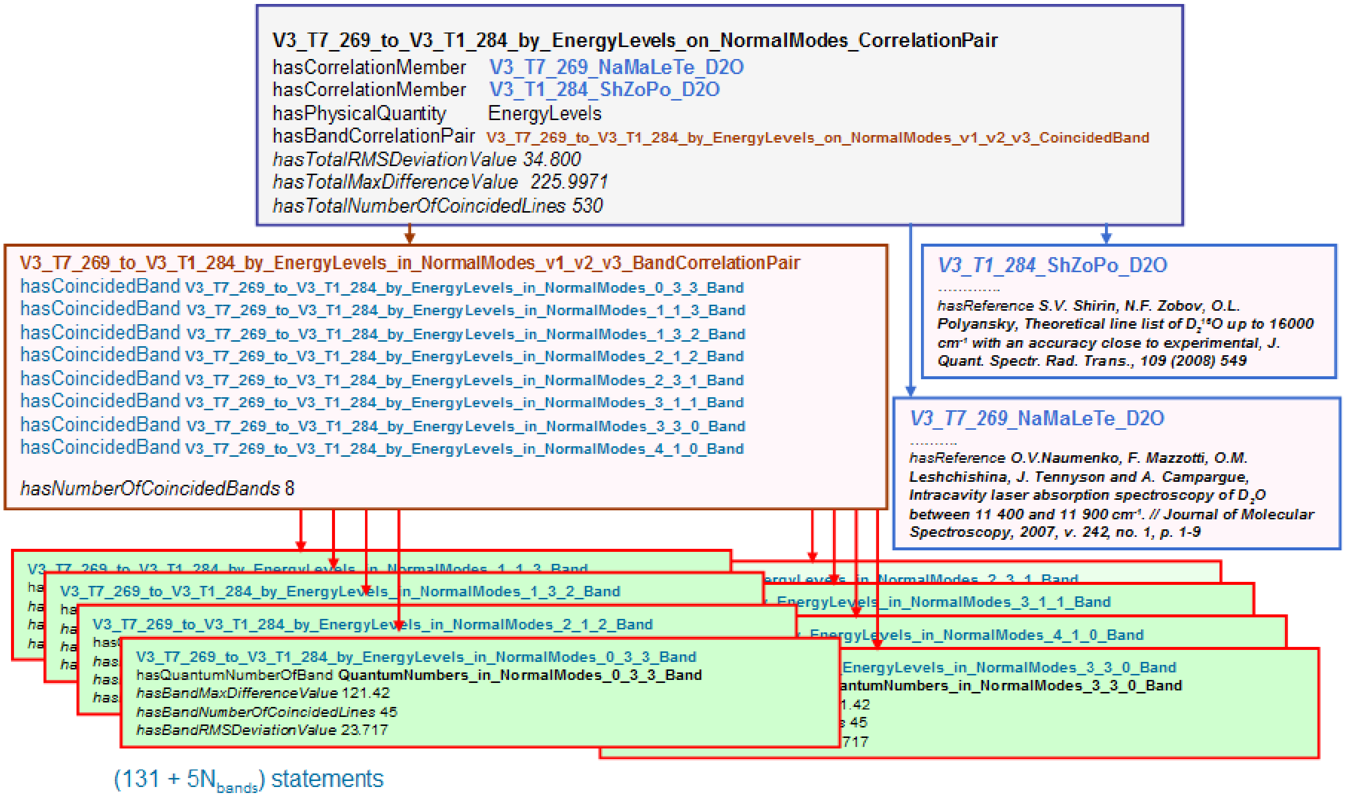
V3_T7_269_to_V3_T1_284_by_EnergyLevels_on_NormalModes_CorrelationPair
hasCorrelationMember V3_T7_269_NaMaLeTe_D2O
hasCorrelationMember V3_T1_284_ShZoPo_D2O
hasPhysicalQuantity EnergyLevels
hasBandCorrelationPair V3_T7_269_to_V3_T1_284_by_EnergyLevels_on_NormalModes_v1_v2_v3_CoincidedBand
hasTotalRMSDeviationValue 34.800
hasTotalMaxDifferenceValue 225.9971
hasTotalNumberOfCoincidedLines 530
V3_T7_269_to_V3_T1_284_by_EnergyLevels_in_NormalModes_v1_v2_v3_BandCorrelationPair
hasCoincidedBand V3_T7_269_to_V3_T1_284_by_EnergyLevels_in_NormalModes_0_3_3_Band
hasCoincidedBand V3_T7_269_to_V3_T1_284_by_EnergyLevels_in_NormalModes_1_1_3_Band
hasCoincidedBand V3_T7_269_to_V3_T1_284_by_EnergyLevels_in_NormalModes_1_3_2_Band
hasCoincidedBand V3_T7_269_to_V3_T1_284_by_EnergyLevels_in_NormalModes_2_1_2_Band
hasCoincidedBand V3_T7_269_to_V3_T1_284_by_EnergyLevels_in_NormalModes_2_3_1_Band
hasCoincidedBand V3_T7_269_to_V3_T1_284_by_EnergyLevels_in_NormalModes_3_1_1_Band
hasCoincidedBand V3_T7_269_to_V3_T1_284_by_EnergyLevels_in_NormalModes_3_3_0_Band
hasCoincidedBand V3_T7_269_to_V3_T1_284_by_EnergyLevels_in_NormalModes_4_1_0_Band
hasNumberOfCoincidedBands 8
V3_T1_284_ShZoPo_D2O
....
hasReference S.V. Shirin, N.F. Zobov, O.L. Polyansky, Theoretical line list of D$_2$^{16}O up to 16000 cm^{-1} with an accuracy close to experimental, J. Quant. Spectr. Rad. Trans., 109 (2008) 549
V3_T7_269_NaMaLeTe_D2O
....
hasReference O.V.Naumenko, F. Mazzotti, O.M. Leshchishina, J. Tennyson and A. Campargue, Intracavity laser absorption spectroscopy of D$_2$O between 11 400 and 11 900 cm^{-1}. // Journal of Molecular Spectroscopy, 2007, v. 242, no. 1, p. 1-9
V3_T7_269_to_V3_T1_284_by_EnergyLevels_in_NormalModes_1_1_3_Band
V3_T7_269_to_V3_T1_284_by_EnergyLevels_in_NormalModes_1_3_2_Band
V3_T7_269_to_V3_T1_284_by_EnergyLevels_in_NormalModes_2_1_2_Band
V3_T7_269_to_V3_T1_284_by_EnergyLevels_in_NormalModes_0_3_3_Band
hasQuantumNumberOfBand QuantumNumbers_in_NormalModes_0_3_3_Band
hasBandMaxDifferenceValue 121.42
hasBandNumberOfCoincidedLines 45
hasBandRMSDeviationValue 23.717
EnergyLevels_in_NormalModes_2_3_1_Band
EnergyLevels_in_NormalModes_3_1_1_Band
EnergyLevels_in_NormalModes_4_1_0_Band
EnergyLevels_in_NormalModes_3_3_0_Band
antumNumbers_in_NormalModes_3_3_0_Band
1.42
45
717
(131 + 5N$_{bands}$) statements

Fig. 3. A structure of an individual which is a formal annotation for T1 and T7 [8] task solutions compared at root-mean-square deviations.

IV. PROBLEMS SOLVED ON BASE OF ONTOLOGY

In our applied ontology, the more complicated constructions of queries form the tasks of calculable systematization of the properties of the information resources. Among such tasks are checks of validity according to appropriate criteria. The precise definition of transition frequencies, intensities, spectral lines etc… is one of the validity criterions, along with restrictions on values for quantum numbers. For example, if the measured transitions with the same assigned quantum numbers have frequencies which differ by less than the measurement error, then they are indistinguishable. Let f_1 and f_2 be measured and calculated frequencies, and a be the experiment error. The condition $|f_1 - f_2| < a$ characterizes the criterion of such a comparison. Its main difficulty is that a is a function of transition frequency. The value of a may differ by orders of magnitude for different frequency intervals and generally decreases as experimental equipment is modernized (upgraded).

Spectroscopic information is systematized according to quantitative values. For example, the definition of measurement precision can lead to different qualitative logical inferences depending on the quantitative value selected by a user. For example, the efficiency of equipment developed for remote sensing in astronomy or planetary atmospheres based on atomic and/or molecular spectra can be overestimated or underestimated. Note that such information resources systematization is necessary for numerous quantitative scientific studies.

Scientific annotations gathered in W@DIS information system form a set of statements forming in its turn an A-box of knowledge bases on water spectroscopy tasks' solutions properties. The properties characterizing the resources form an R-box as well as the classes that unite individuals in T-box. An outstanding feature of a created knowledge base is the possibility to automatically unite some structurally heterogeneous scientific annotations related to the solution of homogeneous spectroscopic tasks.

The task on resources systematization according to validity criteria and the values of their root-mean-square deviations is also solved in the knowledge base of W@DIS (http://wadis.saga.iao.ru) IS along with the other tasks related to scientific annotations.

Table 1 illustrates the results of valid information sources selection from about 800 scientific annotations. The validity was checked according to the restrictions on the values of the quantum number followed from the rules of selection for transitions and restrictions on rotational quantum numbers for energy levels. The asterisk symbol indicates that an annotation for a molecule was modified according to the comments of experts on quantum number assignment correction. As we can see from this table the solutions of T1 and T2 tasks contain a bigger percentage of data arrays that in their turn contain the solutions with incorrect quantum numbers.

Since the direct tasks T1 and T2 describe the states and transitions in the framework of basic principles of quantum physics the system of quantum numbers assignment (normal modes) to states and transitions in water molecule established during the last 70 years can not adequately and unequivocally describe stationary states of a water molecule.

Table 1. The results of selection rules verification in primary data sources [8]

Task /Molecule	T1	T7	T2	T6
H_2O	9 (2)	30 (24)	5 (0)	91 (47)
$H_2{}^{17}O$*	4 (0)	19 (15)	5 (1)	40 (31)
$H_2{}^{18}O$*	4 (0)	18 (18)	5 (1)	59 (35)
HDO	1 (0)	32 (28)	3 (0)	83 (56)
$HD^{17}O$	-	3 (3)	2 (0)	3 (3)
$HD^{18}O$	-	5 (4)	2 (0)	6 (6)
D_2O	1 (1)	18 (8)	3 (0)	38 (26)
$D_2{}^{17}O$			1 (0)	3 (3)
$D_2{}^{18}O$			2 (0)	3 (3)
Total	15 (3)	125 (100)	28 (2)	318 (207)

The other group of tasks is related to checking water spectra; these tasks are in their turn related to criteria governing the existence of these solutions. In our interpretation such criterion corresponds to establishing whether all parts of the checked solution correspond to the published primary data or not. In water spectroscopy the work on collecting and checking the completeness of such a set of data was conducted by the IUPAC group. The software for information resources decomposition was realized in the W@DIS information system which allows to decompose composite sources (for example, Hitran [14] or GEISA [15]) in several minutes. The decomposed result can be interpreted in the following way: if as a result of manipulation a published solution does not contain any parts of the previously published tasks than this solution is fully original. The result of decomposition should be presented in a form of an individual from an ontology included in the scientific annotation. The decomposed result can be interpreted in the following way: if the published solution does not contain any parts of the previously published problems than this solution is fully original. If a published solution contains only the earlier published results than it is either unoriginal or it had a purpose to specify the earlier obtained solutions. Finally, dealing with composite information resources we can single out the unpublished part which obliges a user to thoroughly check the remaining part or to change or delete the resource as an unreliable one. The formation of primary information resources in domains in the framework of the given approach becomes an urgent task, the solution of which would allow us to solve the task of checking the validity of the domain resources in the Internet.

V. EXAMPLES OF CLASSES

The reasoning over an ontology is used to answer queries about its elements. Typical queries include the subsumption between classes (whether a class C subsumes a class D) and the instance checking (whether an individual x is an instance of a class C). In order to answer these queries an inference engine (reasoner) builds a taxonomy for all the named entities in an ontology according to the subsumption relation. This gives a user an option to have the most specific superclasses for any given classes and the most specific types for every individual.

In order to have an easy access to the reasoning results, an ontology usually contains special classes that define the desired query. Canonic data source have data which have to satisfy the physical quantity values restrictions in the domain. In water spectroscopy the constraints are described by the selection rules. These rules are determined by the mathematical model used for the description of the molecules, absorption and emission. The violation of these rules leads to incorrect description of the states or transitions. Quantum numbers play the key role in these rules. The mathematical models of the molecules may be different, so there are different sets of quantum numbers. A few sets of data in water spectroscopy have both sets of quantum numbers (in water spectroscopy, Normal Modes and BT2 notation of quantum numbers are used). Scientific annotation is represented as an individual of an applied ontology and has a group of the properties that characterize the kinds of the selection rules violations.

In applied ontology the definition of the canonic data source related to water molecule may be formulated as:

DataSource **and** hasSubstance **value** H2O **and**
hasOutputData_MD **some** ((hasQuantumNumbers_MD
some (hasQuantumNumbersType **value** NormalModes **and**
hasNumberOfInvalidComparedWithBT2QuantumNumbers
value "0") **and** hasQuantumNumbers_MD **some**
(hasQuantumNumbersType **value** BT2)) **or**
(hasQuantumNumbers_MD **some**
((hasQuantumNumbersType **value** BT2 **and**
hasNumberOfInvalidQuantumNumbers **value** "0") **or**
(hasQuantumNumbersType **value** NormalModes **and**
hasNumberOfInvalidQuantumNumbers **value** "0"))))

This definition includes the case when two different notations are used for the description of data. The error in quantum numbers in any one of two notations has to be interpreted as invalid data. This definition can be easy modified for the other molecules described by normal modes by replacing the names of quantum numbers types or molecules.

VI. CONCLUSION

The model of information resources scientific annotation composing of two parts and represented in an information system in a form of an applied ontology individual used in water spectroscopy allows one to solve the tasks on information resources validity check according to the criteria related to the restrictions on values and existence.

Its representation in a form of an OWL-ontology individual allows us to solve the task on the integration of heterogeneous annotations created by the scientific community for the description of information resources. The scientific annotations applied ontology developed for water spectroscopy is based on the facts that present the properties of the most significant solutions of direct and inverse tasks of water spectroscopy and its isotopomers published in the Internet and in press during the last 60 years.

The scientific annotations for $H_2{}^{17}O$ and $H_2{}^{18}O$ isotopomers description were corrected taking into account the results of paper [5] that allows us to specify the properties of the solution of a series of tasks according the experts' comments. We revealed that the quantitative metadata presented in a form of applied ontology individuals and specified by OWL language allows one to solve the task on automatic definition of validity of primary information sources on spectroscopy tasks' solutions in a distributed information system even if the individuals and taxonomies have different structures.

The further investigation of logical theory of scientific annotations in molecular spectroscopy implies the realization of the following two tasks, namely the task of spectroscopic tasks' solutions automatic classification and the task of checking the existence criterion of the information resources checked for validity.

At present, the solution of the task of spectroscopic tasks' solutions automatic systematization according to the values of root-mean-square deviations is in its completion stage – i.e. the stage of the creation of a user interface for the work with the corresponding part of the knowledge base.

The second task related to the check of the existence restrictions is solved for the case when the decomposition is implemented for the first sources situated at a single node of the distributed information system. More general case requires the solution of the task on the integration of heterogeneous scientific annotations as well as the task of deciding which node of the distributed system would conduct the decomposition, and other tasks.

REFERENCES

[1] The Dublin Core Metadata Initiative, http://dublincore.org/.

[2] J.Tennyson, P.F.Bernath, L.R.Brown, *et al*, IUPAC Critical Evaluation of the Rotational-Vibrational Spectra of Water Vapor. Part I. Energy Levels and Transition Wavenumbers for $H_2{}^{17}O$ and $H_2{}^{18}O$, J. Quant. Spectr. Rad. Transfer, 2009, v. 110, Pages 573-596.

[3] IUPAC project N 2004-035-1-100 «A database of water transitions from experiment and theory». http://www.iupac.org/web/ins/2004-035-1-100.

[4] T. Furtenbacher, A.G. Csaszar and J. Tennyson, MARVEL: measured active rotational-vibrational energy levels, J. Molec. Spectrosc., v. 245, 2007, p. 115-125.

[5] Bykov A.D., Fazliev A.Z., Filippov N.N., Kozodoev A.V., Privezentsev A.I., Sinitsa L.N., Tonkov M.V., Tretyakov M.Yu. Distributed information system on atmospheric spectroscopy // Geophysical Research Abstracts, SRef-ID: 1607-7962/gra/EGU2007-A-01906. 2007. v. 9. p. 01906

[6] Distributed Information System "Molecular Spectroscopy", RFBR grant No. 05-07-90196, http://saga.iao.ru, http://atmos.molsp.phys.spbu.ru/, http://atmos.appl.sci-nnov.ru/.

[7] De Roure D., Jennings N., Shadbolt N. A Future *e*-Science Infrastructure // Report commissioned for EPSRC/DTI Core e-Science Programme. 2001. 78 p.

[8] Privesentsev A.I., Ontological knowledge base implementation and software for information resources description in molecular spectroscopy, Tomsk State University, PhD Thesis, 2009, 238 Pages. Web Information System on Molecular Spectroscopy based on the Knowledge, RFBR grant No. 08-07-00318-a, http://wadis.saga.iao.ru/.

[9] Dmitry Tsarkov and Ian Horrocks. *FaCT++ Description Logic Reasoner: System Description.* In Proc. of the Int. Joint Conf. on Automated Reasoning (IJCAR 2006), volume 4130 of Lecture Notes in Artificial Intelligence, pages 292-297. Springer, 2006.

[10] Dmitry Tsarkov, Ian Horrocks, and Peter F. Patel-Schneider. *Optimizing Terminological Reasoning for Expressive Description Logics.* J. of Automated Reasoning, 39(3):277-316, 2007.

[11] *Sowa J.F.* Knowledge Representation: Logical, Philosophical, and Computational Foundations, Brooks Cole Publishing Co., Pacific Grove, CA, 2000. 594 p.

[12] RFC 2396, Uniform Resource Identifiers, http://www.i-etf.org/rfc/rfc2396.txt

[13] L.S. Rothman, I.E. Gordon, A. Barbe, *et al*,
The HITRAN 2008 molecular spectroscopic database,
J. Quant. Spectr. Rad. Transfer, 2009, v. 110, p. 533-572.

[14] N. Jacquinet-Husson, N.A. Scott, A. Chédin, L. *et al,* The GEISA spectroscopic database: Current and future archive for Earth and planetary atmosphere studies, J. Quant. Spectr. Rad. Transfer, v.109, 2008, p. 1043-1059.