

Challenges in Creating Online Exercises and Exams in Organic Chemistry

Bernhard Jaun and Carlo Thilgen*

Abstract: e-Learning has become increasingly important in chemical education and online exams can be an attractive alternative to traditional exams written on paper, particularly in classes with a large number of students. Ten years ago, we began to set up an e-course complementing our lecture courses *Organic Chemistry I* and *II* within the open-source e-learning environment *Moodle*. In this article, we retrace a number of decisions we took over time, thereby illustrating the challenges one faces when creating online exercises and exams in (organic) chemistry. Special emphasis is put on the development of MOSFECCS (MOlecular Structural Formula Editor and Calculator of Canonical SMILES), our new editor for drawing structural formulae and converting them to alphanumeric SMILES codes that can be submitted as answers to e-problems. Convinced that the possibility for structure input is essential to set up sensible chemistry quizzes and exams, and realising that existing tools present major flaws in an educational context, we decided to embark on the implementation of MOSFECCS which takes into account a number of didactic aspects.

Keywords: e-Learning · Moodle · Online exercises and exams · SMILES generator · Structural formula editor



Bernhard Jaun is honorary professor (emeritus) in the Laboratory of Organic Chemistry, ETH Zürich. After studies in chemistry at ETH Lausanne and ETH Zurich, Bernhard Jaun obtained his PhD from ETH Zurich working under the supervision of A. Eschenmoser and J.F.M. Oth. After a two-year stay as a postdoctoral fellow in R. Breslow's group at Columbia University, New York, he returned to ETH Zurich where he started his own research group in 1989. He was head of the NMR-facility of the laboratory for organic chemistry from 1983 to 2013. Bernhard Jaun's research focus was in the area of physical organic chemistry

and its application to problems relevant in biology. He is a recipient of the Silver Medal of ETH Zurich and the Werner Prize of the Swiss Chemical Society. He taught courses in 1st year organic chemistry (1991–2012) for students of Biology, Pharmaceutical Sciences and Health Sciences and Technology, as well as advanced NMR and advanced physical organic chemistry for students in Chemistry and Interdisciplinary Natural Sciences.



Carlo Thilgen received his diploma degree in Chemistry and his PhD degree from the *Rheinische Friedrich-Wilhelms-Universität Bonn*, working under the supervision of Fritz Vögtle. After a postdoctoral stay in the group of François Diederich at the University of California, Los Angeles, he moved with this group to ETH Zurich, where he became Senior Assistant, Lecturer and, in 2011, Adjunct Professor (*Titularprofessor*). He carried out research in the fields of molecular recognition and fullerenes, in particular the isolation and the chemistry of new higher fullerenes as well as the chirality of pristine

fullerenes and fullerene derivatives. More recently, his interest has focused on the generation of nanostructured surfaces by assembly of porous molecular networks, the ordering of which is based on supramolecular synthons. Carlo Thilgen taught an advanced course in Supramolecular Chemistry next to basic Organic Chemistry courses for students in Chemistry and Interdisciplinary Natural Sciences. Currently, he is teaching first-year Organic Chemistry for students in Biology, Pharmaceutical Sciences and Health Sciences and Technology. In addition, he directs basic and advanced practical Organic Chemistry courses for students in Biology and Pharmaceutical Sciences.

1. Strengthening a Basic Organic Chemistry Course by Online Exercises: Initiation of Our Project

In a basic one-year organic chemistry course at the university, the students are expected to:

- learn a new 'language' (proper terminology) and a new 'script' (structural formulae) to describe the structure and behaviour of molecules;
- train their perception of 3D molecular structure and deduce it from 2D structural formulae; interconvert various 2D representations of 3D structures (perspective drawings, Natta-, sawhorse-, Newman, Fischer projections);
- apply rules and conventions, e.g. systematic nomenclature, stereodescriptors, or Fischer projections;

*Correspondence: Prof. Dr. C. Thilgen
Laboratorium für Organische Chemie
ETH Zurich
Vladimir-Prelog-Weg 3
CH-8093 Zurich
E-mail: thilgen@org.chem.ethz.ch

- carry out a qualitative conformational analysis of basic systems, *e.g.* cyclohexane;
- describe the electronic delocalization in π -systems by use of resonance structures;
- recognize typical structural sub-units (functional groups) which confer distinctive physical or chemical (acidity, reactivity) properties to molecules;
- familiarize themselves with a number of basic organic reactions and apply them to synthetic problems.

Whereas the necessary concepts and theoretical background to achieve these goals are usually imparted in a lecture (2–4 h per week) or can be learnt from textbooks, the acquirement and mastering of the mentioned skills is a matter of practice. At the Department of Chemistry and Applied Biosciences of the ETH Zurich, this is traditionally done in a weekly problem-solving session (1–2 h), in which small groups of students coached by a more advanced fellow discuss the solutions of assignments distributed beforehand as homework.

However, increasing student numbers and the policy to group students from different study programs for introductory courses in basic science, taught as subsidiary subjects, has led to huge classes of up to 500 students. This has made it increasingly difficult to maintain a culture of discussion and dialogue in the exercise groups where attendance has reached a level of about 25 students.

Furthermore, we feel that first-year students, who come from different schools with different core subjects, who have chosen different major subjects at the university and head toward a variety of specializations, have extremely different backgrounds and interests and, consequently, need different extents of training to be enabled to pass a uniform exam at the end of the first year. Besides, we noticed that there is a need for additional practice outside of the lecture period, in particular during the summer break when the students prepare for the exam.

These considerations – but also the curiosity about the feasibility of such an approach – prompted us to develop a series of online trainings in the context of the *Organic Chemistry I* and *II* courses (1st and 2nd semester) for the students of biology, pharmaceutical sciences, and health sciences and technology. They were meant as supplementary training, not as a substitute for the regular assignments discussed in the weekly problem-solving session. The ludic, quiz-type character that online problems can be given in an appropriate e-learning environment, in combination with the recent mushrooming dissemination of tablet computers, makes them attractive brain-teasers available anywhere,

any time. Last but not least, we had the idea at the back of our mind to eventually extend the project toward the elaboration of online exams. With about 500 candidates per year, computer exams appear as an appealing alternative to the traditional manually corrected exams written on paper.

The project was initiated in the fall semester 2007 by one of us (B. J.) with a relatively small set of online exercises embedded within the e-learning environment *Moodle*. In 2008, a FILEP grant ('Fonds zur Finanzierung lehrbezogener Projekte') from the rector of the ETH Zurich allowed a nine-month dedicated full-time employment of a chemist, thereby considerably increasing the number and quality of the quizzes (for an example, see Fig. 5). In 2012, the class was taken over by C. T. who started to build a pool of synthesis problems for the 2nd semester, in which fragmentary reaction schemes have to be completed with products, reactants or reaction conditions. At the same time, we started to convert old 'paper exams' to online problems in order to test the suitability of various types of questions in view of an online assessment. With the gained experience and free infrastructure capacities for a large online exam having finally become available, we are now all set for the Winter 2018 session.

The first version of the open-source e-learning management software *Moodle* was released in 2002^[1,2] and a distribution has been available at the ETH Zurich for about ten years. Accordingly, many people are familiar with the system and the possibilities it offers. Rather than providing a tour d'horizon of our course, we therefore decided to focus on our strategy to overcome a very serious limitation of the text-based *Moodle* system when it comes to chemistry, *i.e.* the missing possibility for a graphical input of molecular structure, an essential aspect in the context of organic chemistry exercises.

2. A Major Flaw of Most e-Learning Environments: The Missing Possibility for Molecular Structure Input

As nicely illustrated by Fig. 1, chemistry – and organic chemistry in particular – is a very 'visual' science. To convey the essential information about a molecular structure, chemists use structural formulae, *i.e.* two-dimensional drawings that are combinations of graphs, letters, numbers and other symbols, possibly including details on the 3D-arrangement of atoms (stereo-information).

Because structural formulae and reaction schemes are so essential in organic chemistry, students must practice this notation or 'script' from the very beginning. Any reasonable system for online exercises must, therefore, accept structural formulae in the formulation of questions as well as the input of answers. Regarding the former requirement, the 'quiz' modules of most current e-learning environments allow the use of graphic elements as part of the question. By contrast, it is a real challenge to provide a tool that allows students to draw structural formulae as answers and have these checked by the computer for correctness. In our view, such a graphical input tool is key for didactic reasons, as it makes a huge difference as to whether students actively try to draw the correct structure themselves or simply choose one out of a given set (multiple choice). Unfortunately, most e-learning environments allow only multiple-choice type answers or the input of text or numbers (alphanumeric strings) which are easily and reliably checked for correctness.

In contrast to human beings, however, software is quite inept at deciding whether two drawings correspond to the same structure or not.^[3] The conventions for structural formulae are rather loose and allow certain things, considered as redundant, to be left out, *e.g.* hydrogen atoms at carbon

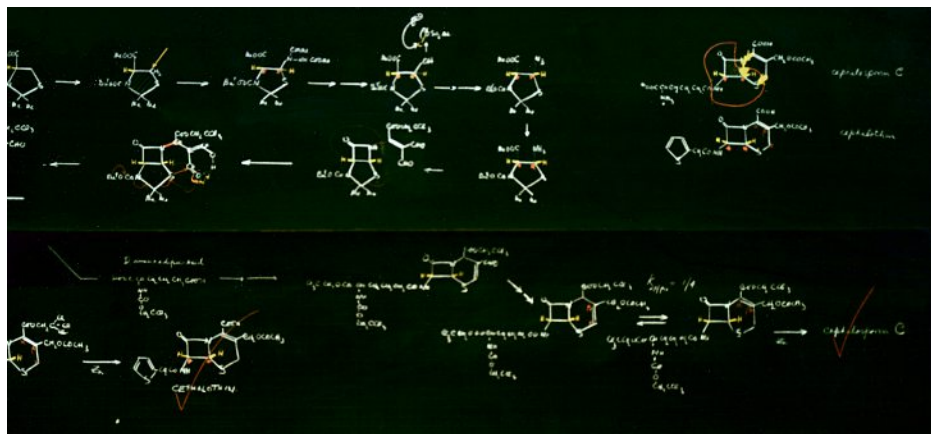


Fig. 1. Chemists communicate through structural formulae. Photograph of the blackboard after a lecture of R. B. Woodward on the total synthesis of Cephalosporin C at ETH Zurich (ca. 1966; Photo: courtesy of Prof. D. Arigoni).

or even at heteroatoms, or to replace the graph of substructures by abbreviations such as R¹, Ph, or by a partial molecular formula such as NO₂, CO₂H, or COOH. Non-bonding electron pairs may be drawn (Lewis structures), but are normally not. In other respects the conventions are very strict: omitting a formal charge on an atom (e.g. H₃C–O instead of H₃C–O[−]) or drawing a filled wedge instead of a dashed one is a fatal error and no longer represents the same molecule.

3. Initial Situation and First Phase of Our e-Learning Project

An e-learning system programmed for online exercises in chemistry ideally includes a tool for drawing structural formulae or whole reaction schemes either directly in the question form or in a separate window. When the drawing is submitted, the corresponding structural information is sent to the server in a suitable format. Server-side software then analyses the answer and returns a feedback (Fig. 2, left).

When we started to think about creating a pool of online quizzes, EPOCH, an e-learning package for organic chemistry coming close to this ideal, had just been created by R. B. Grossman and R. Finkel^[4] in collaboration with ChemAxon, the developers of Marvin.^[5] We had initially hoped to be able to use EPOCH or at least the same standards/interface and tools for our course on the ETH server. Unfortunately, the authors sold the rights to a large publishing company, who renamed the package *ACEorganic* and offered it to the buyers of its text books until 2015. As of late, *ACEorganic* by Grossman and Finkel is again available free of charge for university instructors who wish to author online problems for their students.^[6] However, it can only be used within a cloud computing environment, the servers being located in the US. Only a few of the tools/algorithms used by *ACEorganic* are in the public domain.

Since we had neither the resources, nor the skills, nor the time to program a suitable e-learning environment such as *ACEorganic* from scratch, we decided to adopt a modular approach and use available tools (Fig. 2, right) that would provide the following essential functionalities:

1. An e-learning environment with a built-in quiz module accepting and checking alphanumeric strings as answers.
2. An editor allowing structural formulae to be drawn on a computer or tablet.
3. A tool generating an alphanumeric code that unequivocally reflects the drawn molecular structure and can be copied/pasted into the answer field of the quiz.

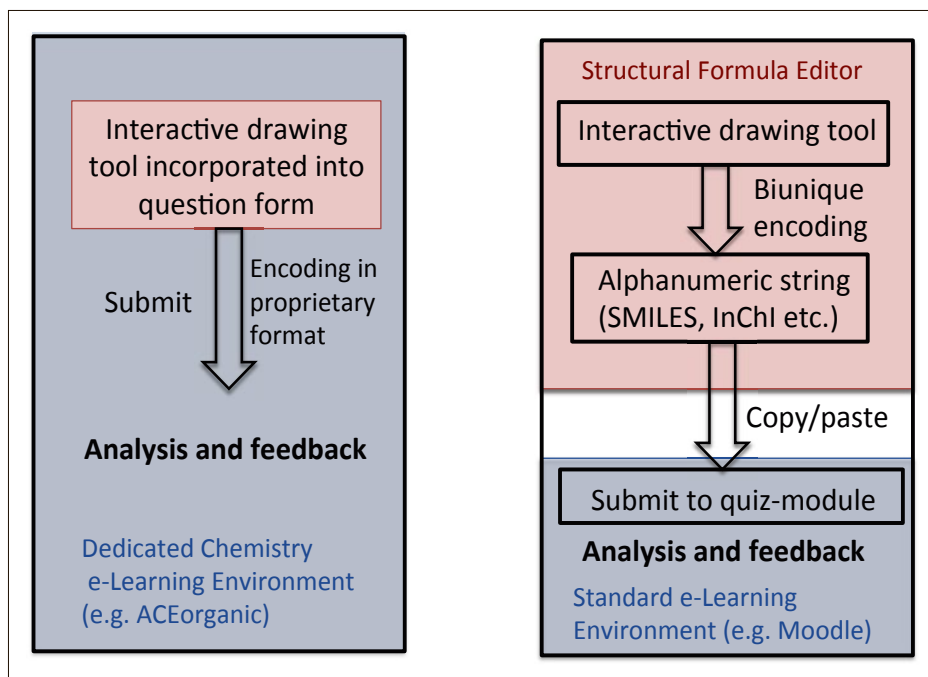


Fig. 2. Two concepts for online exercises in chemistry. Left: integrated e-learning package specifically programmed for chemistry. Right: Modular concept using the quiz module of a standard e-learning package in combination with pre-existing tools for drawing and alphanumeric encoding.

3.1 E-learning Environment

We opted for *Moodle*, one of the e-learning environments that was run and maintained at the time (2006/7) by the teaching support section of the IT services (now LET, *Educational Development and Technology*) at ETH Zurich. Our choice was motivated by the fact that *Moodle* was open source and offered a simple but robust interface already at these early stages. We hoped that a system used and maintained by a broad open community had a better chance of surviving in the rapidly changing IT world than other, mostly commercial, packages – a decision vindicated by subsequent developments. Launched and published in 2002 by Martin Dougiamas,^[1] *Moodle* was designed as a free open source e-learning environment and has since been maintained and further developed by over 600 people from all over the world (our course started with version 1.6, the current version is 3.3) and now counts more than 80⁷000 installations worldwide.^[2]

3.2 Structural Formula Editors and SMILES Generators

We needed a simple but reliable drawing tool that could be invoked *via* a link in the *Moodle* course and would appear in a separate window. Two of the editors available in 2006/7 were considered: JME (Java Molecular Editor) by P. Ertl, Novartis^[7] and *MarvinSketch* by ChemAxon.^[5] As Java Applets, both required the Java runtime environment on the student computers. Neither was open source but both had a documented API (application program-

ming interface). Because JME appeared slightly more stable in our hands, and P. Ertl was kind enough to place the package at our disposal, we opted for JME, which was installed outside the *Moodle* server space but could be called from within *Moodle*. JME had a built-in SMILES-generator (see below), displayed the SMILES code of the drawn molecule(s) in a text window and allowed its transfer to the answer field of *Moodle* quizzes *via* copy/paste. This combination worked well until a few years ago, when increasing concerns about the security of Java forced browser developers and the *Moodle* community to block Java Applets and cross-origin data, and we eventually had to abandon JME. Fortunately, P. Ertl and P. Bienfait released a JavaScript successor (JSME) under the open-source-type BSD (Berkeley Software Distribution) license in time to prevent a forced shutdown of our structure-based online exercises.^[8]

4. The Decision to Develop our own Tool to Draw Structural Formulae and Convert them to Alphanumeric Strings

Under the impression of these events, we nevertheless decided to develop our own structural formula editor, mainly i) to eliminate features that we consider undesirable in an educational context and ii) because of our plans to conduct future 1st year organic chemistry exams online and our uneasiness about using software without knowing its source code.

4.1 Generation of an Alphanumeric Code from a Structural Formula

In 2006/7, the decision to use the SMILES format to transfer structural information from the formula editor to the Moodle quiz answer fields was a practical one: JME produced SMILES codes anyway, and the InChI (International Chemical Identifier),^[9] now the IUPAC-recommended format, was not yet broadly accepted and consists of rather long and incomprehensible strings, even for relatively small molecules. SMILES codes, on the other hand, can be straightforwardly interpreted, *i.e.* re-transformed into a molecular structure by the human mind (Fig. 3).

4.2 SMILES (Simplified Molecular Input Line Entry Specification)^[10]

In 1988, D. Weininger published the rules and algorithms of SMILES, a procedure for encoding the information of a structural formula as an alphanumeric string.^[10b] This encoding, however, is not biunique: A given structural formula can usually be encoded into different formally correct SMILES which, upon back-translation, will all give the same structural formula. The popular program *ChemDraw*,^[11] for example, generates formally correct SMILES but not unique ones and converts formally correct SMILES to a correct structural formula (see Fig. 3). This missing one-to-one correspondence between standard SMILES and structural formulae makes the former unsuitable for database searches. Similarly, using non-biunique SMILES as input in Moodle quizzes would require all possible formally correct SMILES codes for a given structure to be stored as correct answers, which is impossible. To be usable in such a context, the alphanumeric encoding has to be biunique: a given structural formula must be transformed into a single, unique SMILES, regardless of the sequence in which atoms and bonds were drawn. This difficulty, known in graph theory as the ‘canonicalization’ or ‘unique numbering’ problem, was addressed by Weininger *et al.* in a second paper,^[10c] which describes a procedure to generate ‘unique SMILES’. However, others found examples where the procedure of ref. [10c] failed and led to different ‘unique’ SMILES for different drawing sequences.^[12] In the meantime, the rights to SMILES™ had been transferred to *Daylight Inc.* and, although this company claimed to have eliminated the problems, and that the so-called Daylight-SMILES were now truly biunique, neither the corresponding changes in the algorithm nor the actually used code can be peer-reviewed, because they are not in the public domain. This unsatisfactory situation was one of the reasons prompting the NIST (National Institute of Standards and

Technology) and the IUPAC (International Union of Pure and Applied Chemistry) to develop the InChI,^[9] which uses the McKay algorithm^[13] for canonicalization and is biunique. The InChI generation code is open source, but the testing suite for programmers who want to incorporate it into their own software is commercial. The NCI-NIH (National Cancer Institute of the National Institutes of Health) website^[14] offers a tool translating any formally correct SMILES code into the corresponding Daylight-SMILES.

4.3 Editor Design and Features: Conflict between Didactic Goals and Efficiency for Database Searches

JME/JSME, *MarvinSketch* and similar editors are now standard input tools for structure-based database searches. For reasons of convenience, they auto-correct common errors such as pentavalent carbon, automatically generate requisite formal charges and offer palettes with ready-made functional groups *etc.* With the objective of finding a maximum if not all of the entries for a given substance in a database, modern structure editors try to merge different representations such as separate resonance structures and even tautomers. Daylight-SMILES tries to recognize and assign aromaticity for this purpose, whereas InChI only uses connections and does not distinguish between bonds of different bond order.

All these features are inappropriate for a formula editor used to teach first-year organic chemistry, where the formalism and rules of classical structure theory or electron delocalization are key subject matters: the students must be able to make mistakes and learn from them! Online exercises, let alone exams, are practically worthless if the used structure editor auto-corrects errors and gives out warnings if valences are exceeded or formal charges forgotten. What is the point of a tool that transforms all resonance structures or tautomers into a single one or tells the students that a particular cycle is aromatic

when they are supposed to recognize it on their own?

5. MOSFECCS – A Newly Developed Structural Formula Editor and SMILES Generator

The drawing tool and SMILES generator MOSFECCS (MOlecular Structural Formula Editor and Calculator of Canonical Smiles, Fig. 4), which we introduced in 2016, follows a more didactic perspective, tolerating many kinds of mistakes and treating each resonance structure or tautomer as a distinctive species.

MOSFECCS generates biunique SMILES codes that are, as far as tested, identical to Daylight-SMILES produced by the NCI-NIH translator, except in cases where we deliberately deviate from the Daylight-SMILES convention. This applies to rings with a maximum number of conjugated double bonds: Instead of trying to assign aromaticity (lower-case letters are used by Daylight-SMILES for atoms in *aromatic* rings), MOSFECCS uses small letters for all atoms of even-membered *mancude* rings (see Fig. 4).^[15] This has nothing to do with the concept of aromaticity but simply avoids having to program all their trivial resonance or double-bond-shifted structures as correct answers in a quiz (each unsymmetrically substituted benzene ring in a molecule would multiply the number of required entries by two). For the purposes relevant in a basic organic chemistry course, there are no negative consequences if the SMILES code ignores the position of alternating double bonds in even-membered mancude rings. For other unsaturated even-membered rings (*e.g.* σ -complexes) and for odd-membered unsaturated rings (*e.g.* pyrrole, cyclopentadienyl anion, tropylium cation), the generated SMILES code reflects only the particular resonance structure drawn.

Stereochemistry is a pivotal subject in the first year of an organic chemistry course. Following the so-called isomeric-SMILES

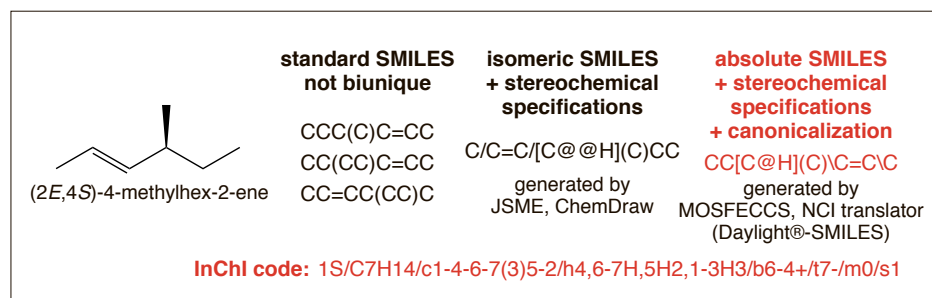
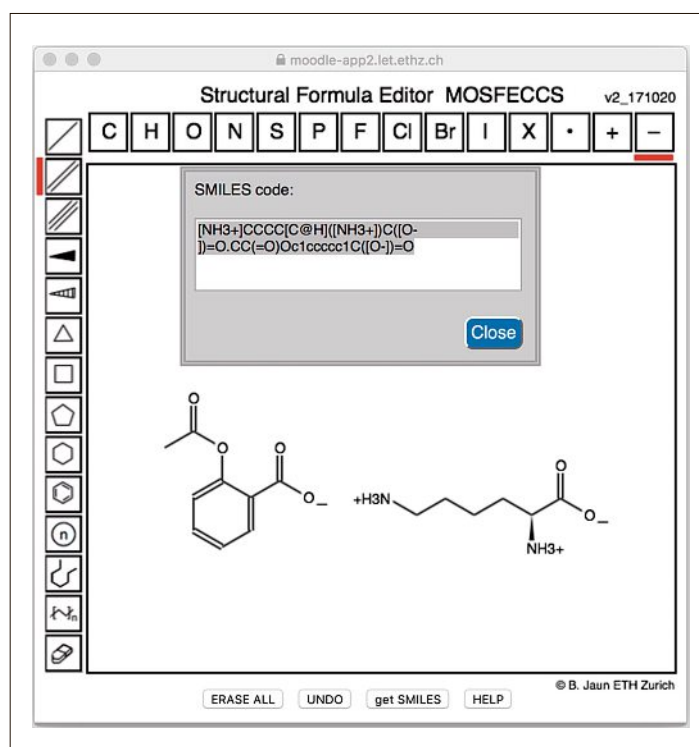


Fig. 3. Encoding of structural formulae into alphanumeric strings. Different SMILES conventions, the InChI code, and output of SMILES generators used by different structural formula editors with (2E,4S)-4-methylhex-2-ene as example. To be useful for online quizzes and tests, the encoding must be biunique (red).

Fig. 4. MOSFECCS (MOlecular Structural Formula Editor and Calculator of Canonical SMILES): View of the canvas with the drawing toolboxes at the upper and left edges. The currently active tools are marked with a red bar. The inset shows the SMILES code '[NH3+]CCCC[C@H]([NH3+])C([O-])=O.CC(=O)Oc1ccccc1C([O-])=O' generated for the drawn formula (salt from acetylsalicylic acid and L-lysine) that can be copied and pasted into the answer field of a quiz (not shown).



(integral component of Daylight-SMILES) convention, MOSFECCS uses the symbols @ and @@ to specify the absolute configuration of stereogenic centres and the marks // and \ for the configuration of double bond units (Fig. 4).^[16] In the case of stereogenic centres with an implicit hydrogen or a lone pair as one of the four ligands, the 3D-coordinates of this 'invisible' ligand have to be deduced from the way the bonds to the three visible ones are drawn. As analysed comprehensively in the technical InChI manual,^[17] only a subset of all possible ways to draw a centre with tetrahedral or trigonal-pyramidal bond geometry defines its absolute configuration unambiguously. If MOSFECCS encounters equivocally drawn bond geometries at stereogenic centres during the SMILES generation, it marks them on the canvas with a red square and issues a warning instead of a SMILES code, thus forcing the student to improve the drawing until it is unambiguous.

When the answer to a quiz encompasses several structures (question type: "draw the three best resonance structures of prop-2-enal"), it is convenient to draw them all at once on the canvas and submit them together to the SMILES generator (Fig. 4). The SMILES convention allows for specification of multiple structures by a single string in which the codes for the individual molecules are simply concatenated with a period as separator. However, the convention includes no rule about the order in which the individual sub-strings should be linked together; with most editors, the sequence depends on the order of drawing or the position of the struc-

tures on the canvas. For online exercises this is inexpedient, because it requires all permutations of individual, dot-separated sub-strings to be programmed as correct answers in the quiz module. To eliminate this nuisance, MOSFECCS always concatenates in a well-defined order by arranging the individual SMILES sub-strings according to decreasing length and – if the length is the same – in alphabetic order (Fig. 4).

The scripting part of MOSFECCS is written entirely in JavaScript (ECMA script 5.1) and uses Canvas elements as defined in HTML 5 for interactive drawing. So far, it has been tested for the current versions of the browsers Firefox (Windows 10, MacOS, iOS, Android), Safari (MacOS, iOS), Edge (Windows 10), Chrome (MacOS, Android) and the Safe Exam Browser^[18] (MacOS and Windows 10). In recent years, more and more students use tablets rather than laptop or desktop computers. Sketching a structural formula with fingertips or a stylus instead of a mouse requires some adjustment for any drawing program. Currently, MOSFECCS runs on iOS-, Android-, and Windows 10-based tablets.

6. A Particular Challenge: The Handling of Partially Correct Answers in an Online Test

6.1 Feedback to Wrong or Partially Correct Answers

Online problems are most valuable if they provide a feedback which is specific to the mistake the student made in the case

of incorrect or partially correct answers (Fig. 5). Human teachers can often recognize the error in reasoning which leads to a particular wrong answer and distinguish it from simple guessing or a slip of the pen. When wrong answers that are related to foreseeable errors in reasoning are submitted in online quizzes, the students can be given specific feedbacks with tips enabling them to take the next hurdle and improve in the next trial. Actually, the teacher can identify the most frequent mistakes by looking at the statistics of given answers and improve the feedback accordingly over time. One problem, however, persists: computers cannot distinguish between faulty reasoning and slips of the pen. Whenever the latter add on top of the former (and this happens more often than one would believe), a useful feedback is not possible.

6.2 Online Exams: The Grading of Partially Correct Answers

For authors of online exams, automatic grading of partially correct answers that include anticipatable mistakes is a particular challenge. In a nomenclature question, for example, one would normally assign a partial grade if the systematic name submitted for a given enantiomer is correct, but the stereodescriptor is either wrong or missing. If the molecule contains a single stereogenic centre only, this situation is easily handled by an additional entry in the list of expected answers. If, however, it includes more than one stereogenic centre, the number of extra entries in the answer list grows very rapidly. One way to reduce the list of answers considered for partial grading is the use of wildcards (similar to UNIX shells, Moodle interprets the asterisk * as ≥ 0 arbitrary characters). The text-based question types of Moodle, i.e. SHORTANSWER and CLOZE, accept wildcards in the programmed answers and correctly match these against the student's submitted answer. Regular expression (*regex*) matching, known from the *grep* command in UNIX shells or from programming languages such as Perl or PHP, would be much more powerful than a simple wildcard. *Regex* modules for Moodle have indeed been developed, but unfortunately they are not part of the Moodle core distribution and have to be installed from the third party plugin repository.^[19] This means that their operability may be lost after the next Moodle upgrade, which is, alas, a strong argument against using *regex* in the course.

The incorporation of wildcards in SMILES strings is of limited usefulness. Due to the process of canonicalization, small changes in a structure, such as the exchange of a methyl for an ethyl substituent or a different substitution pattern, may lead to a completely different sequence of atom

ETH zürich

→ Dashboard → Meine Kurse → 529-1011-00L Organische Chemie I (für Biol./Pharm.... → 5 - Stereochemie
→ Stereochemie - Training 3 (Fischerprojektionen) → Vorschau

Zeichnen Sie von folgender Fischerprojektion eine Keilstrich-Formel.

Tragen Sie den SMILES-Code in das Feld unten ein.

$$\begin{array}{c}
 \text{CO}_2\text{H} \\
 | \\
 \text{HO} - \text{C} - \text{H} \\
 | \\
 \text{HO} - \text{C} - \text{H} \\
 | \\
 \text{H} - \text{C} - \text{OH} \\
 | \\
 \text{HO} - \text{C} - \text{H} \\
 | \\
 \text{CO}_2\text{H}
 \end{array}$$

Antwort:

Prüfen

Specific feedback given after submitting this answer

Die Konfiguration von ZWEI stereogenen Zentren ist leider falsch!

Teilweise richtig

Bewertung für diese Einreichung: 0.50/1.00. Diese Einreichung berücksichtigt einen Abzug von 0.25.

Frage 8

Die Konfiguration von ZWEI stereogenen Zentren ist leider falsch!

Schrittweise Umwandlung der Fischer-Projektion in eine Keilstrichformel mit Zickzack-Rückgrat:

$$\begin{array}{c}
 \text{CO}_2\text{H} \\
 | \\
 \text{HO} - \text{C} - \text{H} \\
 | \\
 \text{HO} - \text{C} - \text{H} \\
 | \\
 \text{H} - \text{C} - \text{OH} \\
 | \\
 \text{HO} - \text{C} - \text{H} \\
 | \\
 \text{CO}_2\text{H}
 \end{array}
 \Rightarrow
 \begin{array}{c}
 \text{CO}_2\text{H} \\
 | \\
 \text{HO} - \text{C} - \text{H} \\
 | \\
 \text{HO} - \text{C} - \text{H} \\
 | \\
 \text{H} - \text{C} - \text{OH} \\
 | \\
 \text{HO} - \text{C} - \text{H} \\
 | \\
 \text{CO}_2\text{H}
 \end{array}
 \Rightarrow
 \begin{array}{c}
 \text{HO}_2\text{C} - \text{C}(\text{OH}) - \text{C}(\text{OH}) - \text{CO}_2\text{H} \\
 | \quad | \\
 \text{H} \quad \text{H} \\
 \text{+ Drehung jeder zweiten CH(OH)-Gr. um } 180^\circ
 \end{array}$$

Die Keilstrichformel kann durch eine 180°-Drehung um verschiedene Achsen verschieden dargestellt werden.

Die richtige Antwort ist: O[C@H]([C@H](O)[C@H](O)C(O)=O)[C@H](O)C(O)

General feedback given after submitting the test

Fig. 5. Sample quiz question from our Moodle Course *Organic Chemistry I*: Transcription of a given Fischer into a Natta projection. The structural formula editor MOSFECCS is used by the student to draw the Natta projection and generate its SMILES code which is copied/pasted into the answer field ('Antwort') (left). In the shown example, the proposed structure and, by implication, the input SMILES code contains two mistakes. After this partially wrong answer is submitted by pressing the check button ('Prüfen'), Moodle returns a specific feedback saying that two stereogenic centres have the wrong configuration. The student can now revise the answer and resubmit. After working through all the questions of a test, the student receives a general feedback including the correct answers as well as any explanations elaborated by the teacher (right).

symbols inside the SMILES string. On the other hand, when the correct answer to a quiz question is a multi-structure SMILES, wildcards do allow to check for the presence of one or several of the constituent sub-codes within the aggregate SMILES string. This allows for automatic grading when a student finds only two out of a total of three isomers or resonance structures.

In view of the above shortcomings, it is indispensable that, after an online exam, the teacher carefully reviews the submitted answers, notably those marked as wrong or partially correct by the computer. Fortunately, Moodle makes this perusal of the students' answers easy and allows for manual modification of the grades. For an efficient and reliable human scrutiny of submissions consisting of SMILES codes, however, these have to be re-transformed into structural formulae. For this purpose, we have developed a script that imports the SMILES answers of all candidates into a spreadsheet. With the ChemDraw plugin for Excel^[20] the codes can then be automatically re-converted to structural formulae for visual inspection.

7. Current Limitations and Possible Extensions of MOSFECCS

The current version of our structural formula editor and SMILES generator (MOSFECCS version 2) is tailored for organic chemistry. To make it useful for inorganic chemistry as well, other types of bonds such as coordinative bonds, π -coordination, *etc.* and stereoisomerism at centres with square-planar, trigonal-bipyramidal and octahedral bond geometry would have to be included (the latter type is already defined in the Daylight-SMILES convention).^[16] For questions involving chemical reaction schemes, it would be desirable to extend the alphanumeric encoding in order to include reaction arrows and specifications for reactant, reagent/catalyst and product, similar to the reaction-SMILES notation published by Daylight.^[21] Questions expecting the students to formulate reaction mechanisms would require the optional drawing of lone pairs and curved arrows in the structural formula editor as well as the encoding of the electron-shift notation in the

alphanumeric code as implemented (via MarvinSketch) in the ACEorganic e-learning environment.

8. Conclusions

After setting up and running a 1st year organic chemistry Moodle course for a decade, we are pleased to see that this supplementary class module is becoming increasingly popular among the students. This is partly due to the fact that, with the advent of tablet computers and smartphones, students are getting used to the pervasive use of electronic tools and resources. But it is also related to the refinement of the course material, in particular the multiplication of the online quizzes which now cover a large range of topics. We learnt the lesson that while an elaborate course with online problems allows a very convenient training of students at almost no cost per capita, the effort and time required for the setup and maintenance are substantial. Since the didactic value of online quizzes crucially depends on the quality of the feedbacks, experienced teachers are needed for the design and elaboration, whereas the technical editing and testing can be delegated to others.

The initial choice of e-learning environment, core software tools, structural formula editor, and the interface to convert the formulae to alphanumeric data has far-reaching consequences and should be well considered. For a course with more than a thousand quiz questions and an even greater number of answers/feedbacks, the decision to switch to other software components or standards requires a substantial part of the work already invested to be redone. For the same reason, it is strongly advised to resist the temptation of using new, fancy question types that offer tools such as an integrated molecular formula editor^[22] or the use of powerful regular expressions. Such tools periodically appear in the Moodle world, but they usually have a limited lifetime because they are not incorporated in the Moodle core modules.

The history of our project also reveals the problems associated with a dependence on proprietary software or on software as a service (SaaS, cf. ACEorganic).^[6] In the world of scientific computing, there are many examples of academic authors having created an ingenious software tool which, at some point, was commercialized without publication of the code. The ensuing trouble is not primarily of pecuniary nature and many copyright owners even allow a free non-commercial use of their software (Marvin, JME, JSME). The main inconvenience is the lack of control over one's most important tools and the non-existent guarantee that proprietary software

will be maintained and available over time. In the context of online chemistry exams, undisclosed source code may lead to an additional complication: What is the proper response to the possible appeal of a student claiming to have drawn the correct molecular structure but that a bug in the formula editor caused an error in the encoding?

Considering these risks, we feel that the only satisfactory solution is the use of open-source software throughout. Accordingly, we plan to make the structural formula editor MOSFECCS available under an open-source license as soon as its development and initial round of testing have been completed.

Overall, our setting up of *Moodle* courses with online exercises, although quite operose, was and continues to be a rewarding and instructive experience. As a result, the students have at their disposal a considerable and ever-growing pool of online quizzes, which they can use *ad libitum*. We – and possibly the students too – are now anxious to take the next step, namely the implementation of online exams including the submission of structural formulae as answers.

Acknowledgements

We gratefully acknowledge Markus Beck for contributing many new online exercises in 2008, the LET team (ETH Zurich) for supporting us from the beginning and maintaining the *Moodle* installation at the school, ETH Zurich for financial support (FILEP grant, 2008) and the many teaching assistants as well as a number of students of the courses *Organic Chemistry I and II for students of Biology, Pharmaceutical Sciences, and Health Sciences and Technology* (ETH Zurich), who contributed ideas or reported bugs, thereby helping to improve the online course. We thank Dr. Peter Ertl (*Novartis*) for letting us use JME and JSME.

Received: October 7, 2017

- [1] M. Dougiamas, P. C. Taylor, 'Interpretive analysis of an internet-based course constructed using a new courseware tool called Moodle', Proceedings of the Higher Education Research and Development Society of Australasia (HERDSA) 2002 Conference, Perth, Western Australia.
- [2] <https://moodle.org>, accessed 29 Sept. 2017.
- [3] For a project to directly interpret bitmap graphics of scanned structural formulae see: I. V. Filippov, M. C. Bicklaus, *J. Chem. Inf. Model.* **2009**, 49, 740.
- [4] R. R. Chamala, R. Ciochina, R. B. Grossman, R. A. Finkel, S. Kannan, P. Ramachandran, *J. Chem. Educ.* **2006**, 83, 164.
- [5] <https://www.chemaxon.com>, accessed 29 Sept. 2017.
- [6] <https://ace.chem.illinois.edu/ace/public/features/index.html>, accessed 29 Sept. 2017.
- [7] P. Ertl, *J. Cheminf.* **2010**, 2, 1; see also: <http://www.molinspiration.com/jme/index.html>, accessed 29 Sept. 2017.
- [8] B. Bienfait, P. Ertl, *J. Cheminf.* **2013**, 5, 24.
- [9] S. R. Heller, A. McNaught, I. Pletnev, S. E. Stein, D. V. Tchekhovskoi, *J. Cheminf.* **2015**, 7, 23.
- [10] Concepts for encoding the information contained in a structural chemical formula as an alphanumeric string – more amenable to electronic data processing than a picture – have a long history, ranging from a) the Wiswesser Line Notation: W. J. Wiswesser, *J. Chem. Inf. Comput. Sci.* **1982**, 22, 88; over b) SMILES (Simplified Molecular Input Line Entry Specification): D. Weininger, *J. Chem. Inf. Comput. Sci.* **1988**, 28, 31; and c) its variants, the isomeric-SMILES, unique-SMILES: D. Weininger, A. Weininger, J. L. Weininger, *J. Chem. Inf. Comput. Sci.* **1989**, 29, 97; d) Daylight-SMILES™: <http://www.daylight.com/smiles/index.html>, accessed 29 Sept. 2017; e) and open-SMILES: <http://opensmiles.org>, accessed 29 Sept. 2017; Specification: <http://opensmiles.org/opensmiles.html>, accessed 29 Sept. 2017; to the InChI (International Chemical Identifier) code: see ref. [9].
- [11] ChemDraw® and ChemOffice® 16, Perkin Elmer Inc.
- [12] G. Neglur, R. L. Grossman, B. Liu, in 'Data Integration in the Life Sciences', Proceedings of the Second International Workshop DILS, **2005**, Eds. B. Ludäscher, L. Raschid, pp 145-157, Springer, **2005**.
- [13] a) B. D. McKay, *Congr. Numer.* **1980**, 30, 45; b) B. D. McKay, A. Piperno, *J. Symb. Comput.* **2014**, 60, 94.
- [14] <https://cactus.nci.nih.gov/translate>, accessed 29 Sept. 2017.
- [15] IUPAC, 'Compendium of Chemical Terminology', 2nd ed. (the 'Gold Book'), compiled by A. D. McNaught and A. Wilkinson, Blackwell Scientific Publications, Oxford, **1997**. XML on-line corrected version: <http://goldbook.iupac.org> (2006-) created by M. Nic, J. Jirat, B. Kosata; updates compiled by A. Jenkins. ISBN 0-9678550-9-8. <https://doi.org/10.1351/goldbook>, accessed 29 Sept. 2017. Last update: 2014-02-24; version: 2.3.3. DOI of this term: <https://doi.org/10.1351/goldbook.M03695>.
- [16] Daylight Theory Manual, Version 4.9, 2011, Daylight Chemical Information Systems, Inc. Laguna Niguel, CA, <http://daylight.com/dayhtml/doc/theory/theory.smiles.html>, accessed 29 Sept. 2017.
- [17] S. E. Stein, S. R. Heller, D. V. Tchekhovskoi, I. Pletnev, IUPAC International Chemical Identifier (InChI) InChI version 1, software version 1.04, **2011**, Technical Manual.
- [18] <http://safeexambrowser.org>, accessed 29 Sept. 2017.
- [19] https://docs.moodle.org/33/en/Regular_Expression_Short-Answer_question_type, accessed 29 Sept. 2017.
- [20] The ChemDraw for Excel plugin (included in Chem Office 16, Perkin Elmer Inc.) is only available for Microsoft Windows and requires the 32bit version of Excel (Microsoft Office).
- [21] A. R. Leach, J. Bradshaw, D. V. S. Green, M. M. Hann, *J. Chem. Inf. Comput. Sci.* **1999**, 39, 1161.
- [22] <https://docs.moodle.org/33/en/question/type/jme>, accessed 29 Sept. 2017.