



Reaction Scheme: Capture, Iteration, and Parsing

Zhenbin (Benjamin) Li, Ph.D.

Research Data Integration and Logistics Service



Boehringer
Ingelheim



Boehringer Ingelheim Center
Our headquarters in Ingelheim, Germany

- Focus on Human Pharmaceuticals and Animal Health
- Family-owned global corporation
- Founded 1885 in Ingelheim, Germany
- Employees worldwide: 41,534
- Total revenue: EUR 12,721 million
- Expenses for R&D: EUR 2,215 million



CROs: India and China

Our R&D Focus on 6 therapeutic areas



Research Areas	Examples of disease areas	Sites
Respiratory diseases	Asthma, COPD, Idiopathic pulmonary fibrosis	Biberach (Germany)
Cardiometabolic diseases	Atherosclerosis, Dyslipidemia, Diabetes, Chronic Kidney Disease	Biberach (Germany) Ridgefield (USA)
Oncology	Lymphomas, Leukaemias, Solid tumours	Vienna (Austria)
Neurological diseases	Alzheimer's disease, Chronic pain, Migraine, Parkinson's disease	Biberach (Germany)
Immunology & Inflammation	Multiple sclerosis, Psoriasis, Rheumatoid arthritis	Ridgefield (USA)
Virology	Hepatitis C, HIV/AIDS	Laval (Canada)



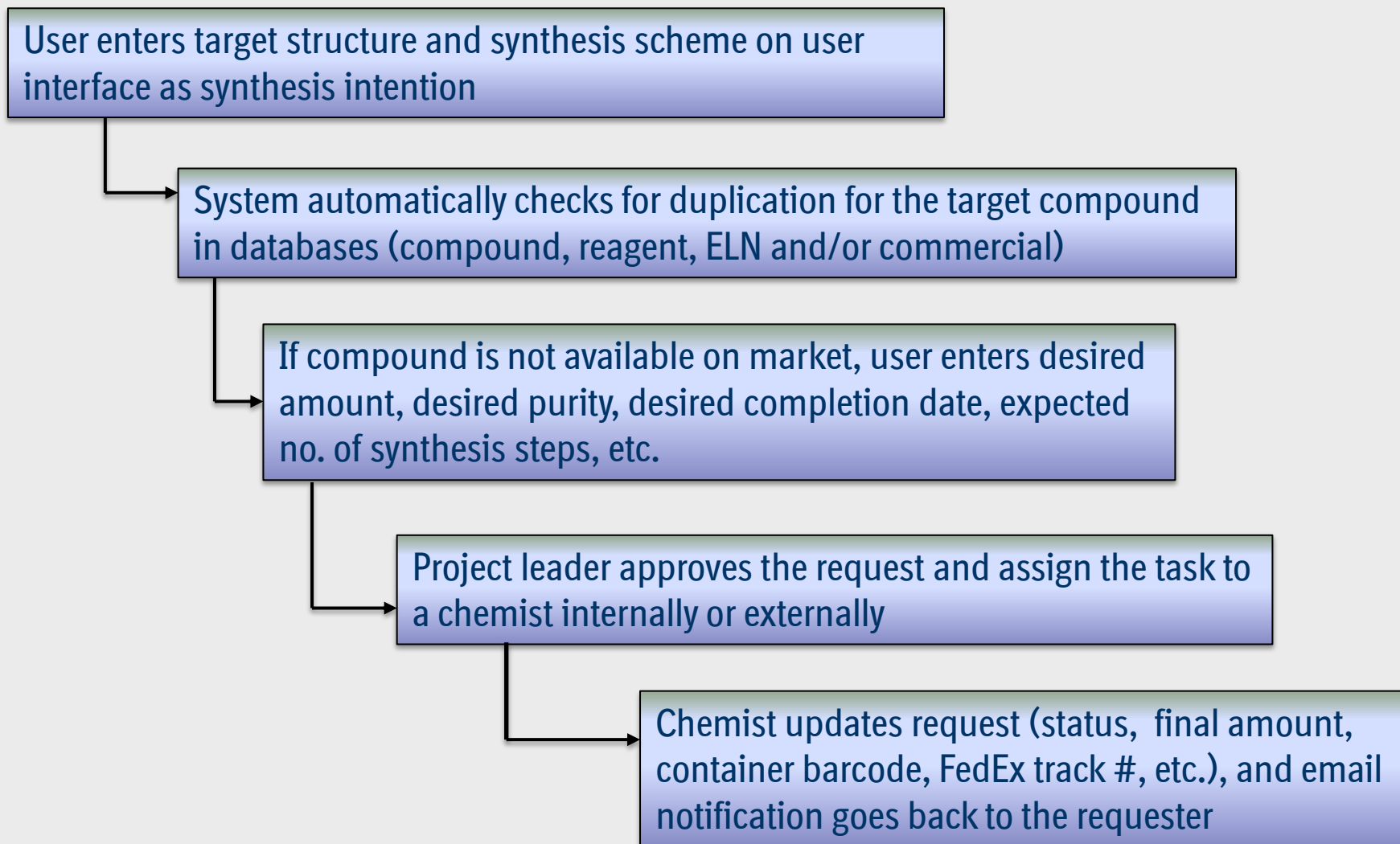
2010, US FDA Approves Pradaxa® (dabigatran etexilate) – a breakthrough treatment for stroke risk reduction in non-valvular atrial fibrillation

Decision marks first approval of a new oral anticoagulant in the USA in 50 years

TRADJENTA



2011, FDA approves Boehringer Ingelheim and Eli Lilly's linagliptin tablets for the treatment of type 2 diabetes alongside diet and exercise measures



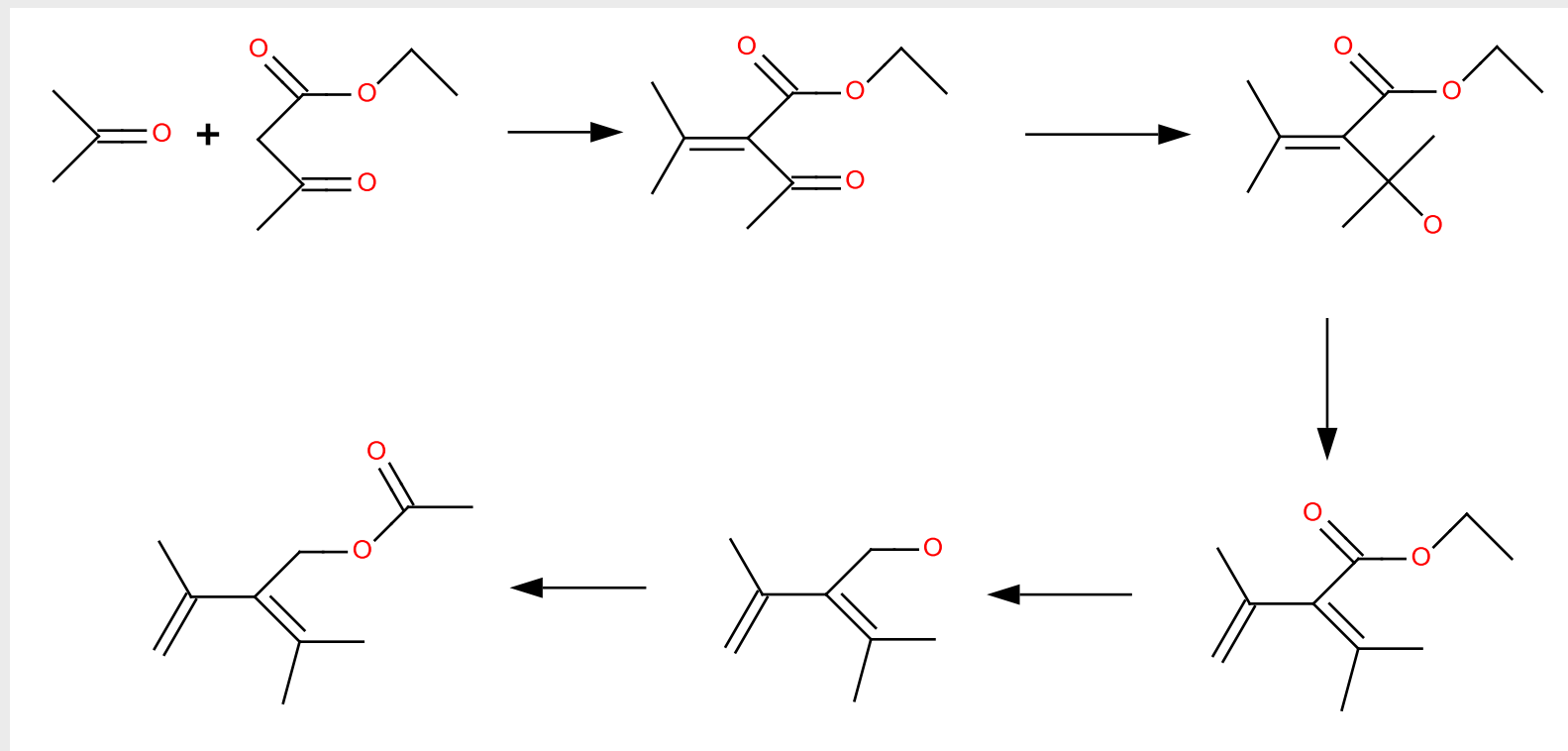
Benefits realized

- Information capture through centralized user interface (request details, structures, procedures, references, and timestamp)
- Redundancy elimination (compound, inventory, ELN and request databases)
- Automatic email notification to the appropriate CRO managers for prompt actions
- Implementation of a structure searching feature
- Improve communication with CROs
- Efficient status tracking
- Automatic registration into reagent database
- Automatic or user-defined metrics calculation and reporting
- Timely management responses and decisions

Challenge

- Compound and logistics are stored in database, except for individual reactions in synthesis scheme

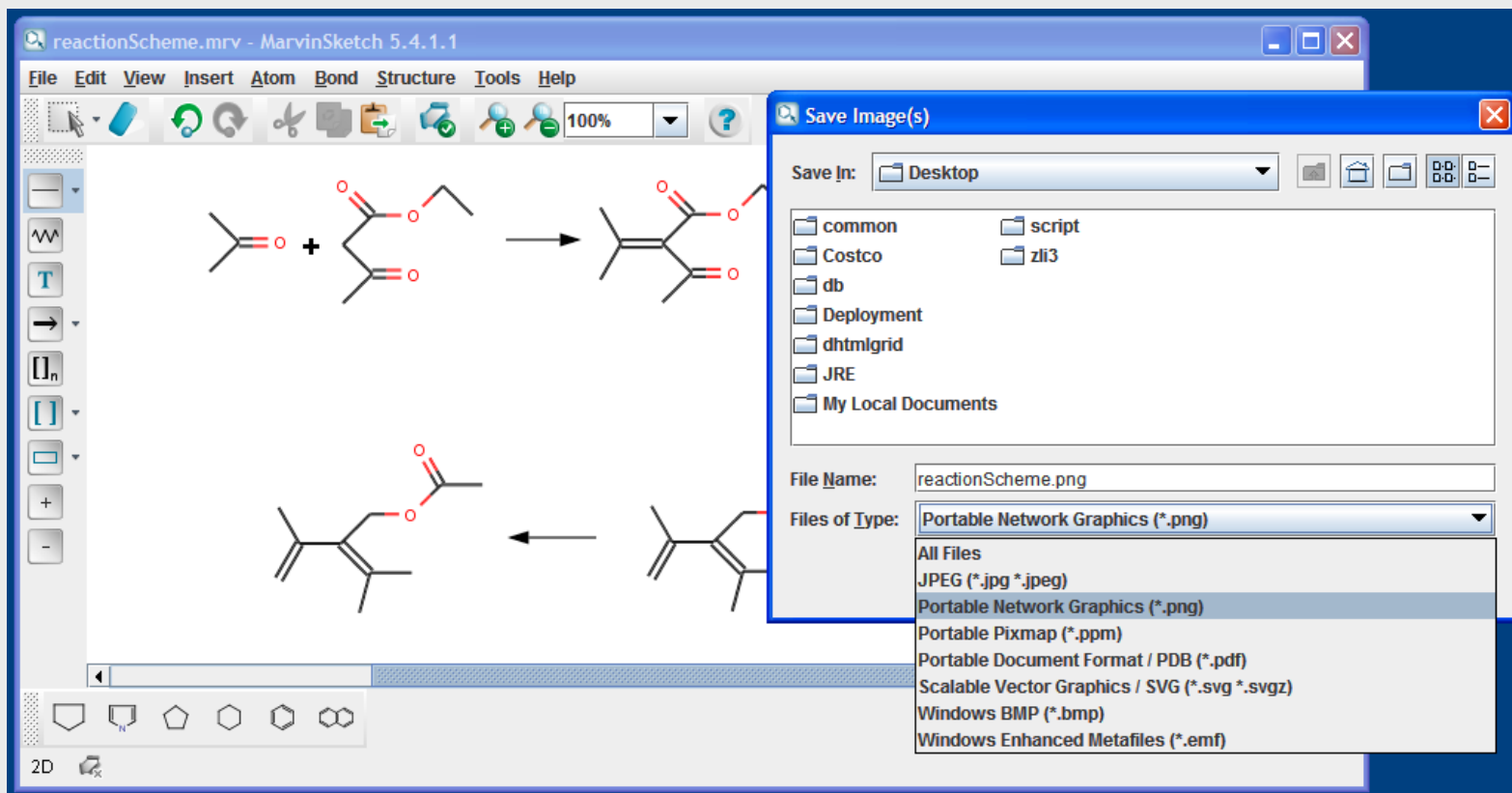
Aldol condensation of acetone and ethyl acetoacetate gave β -keto-ester **3**. A Grignard reaction involving methylmagnesium bromide provided alcohol **4**, which was subjected to acid catalyzed elimination to give diene **5**. Reduction and acylation gave diene **7** (Scheme 3, compound 1).



* http://en.wikipedia.org/wiki/Nicolaou_Taxol_total_synthesis

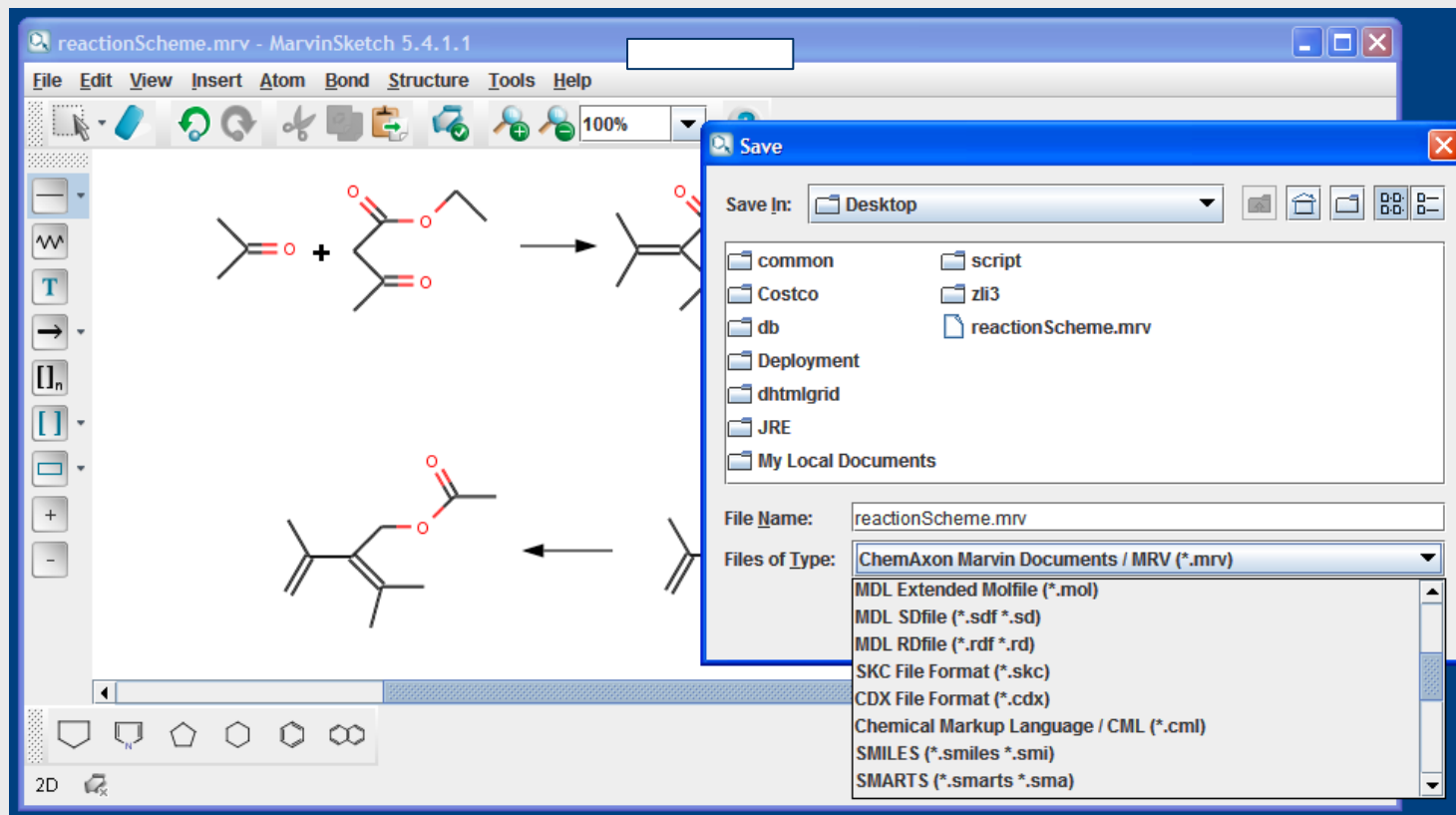
Image File Formats

- MarvinSketch allows saving reaction scheme in image files on the interface.
- Different image formats are available.
- The image file can also be generated programmatically so that the image file can be directly inserted into file storage (database or file folder)

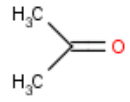
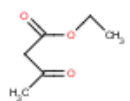
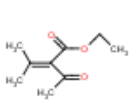
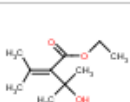
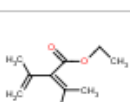
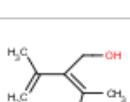
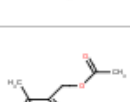


Chemistry File Formats

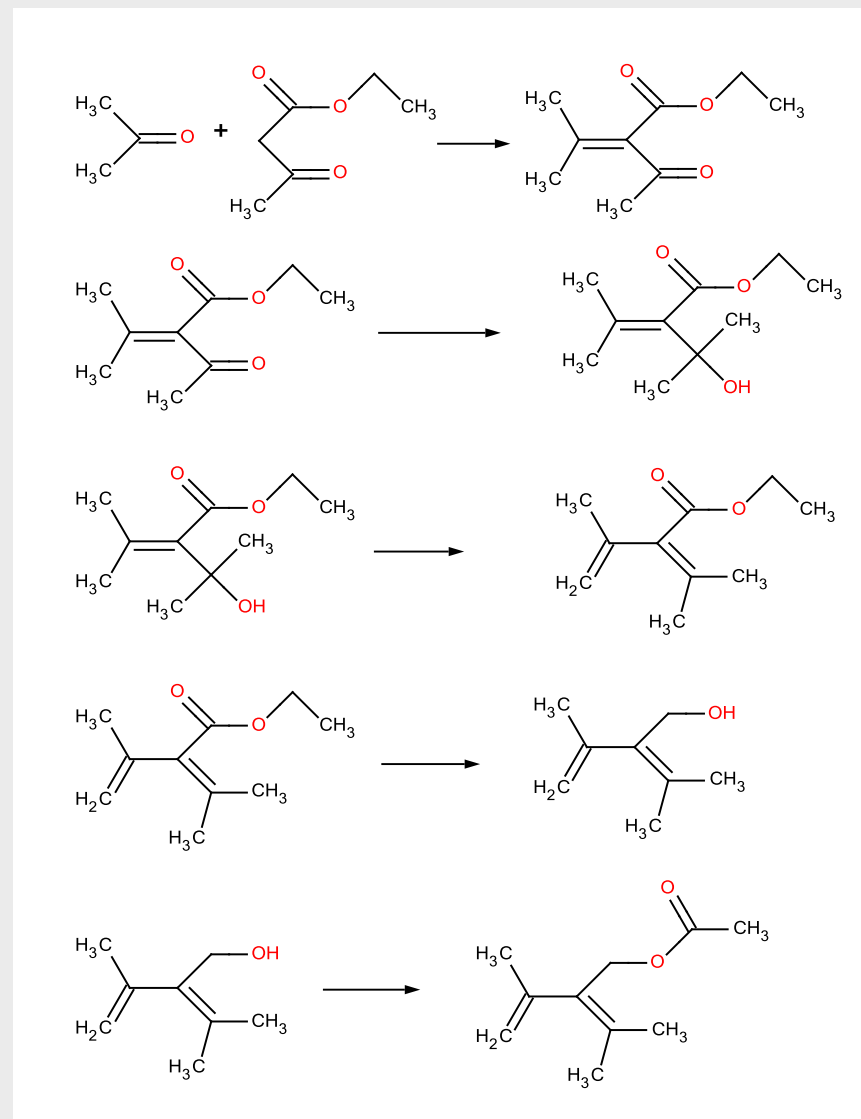
- MarvinSketch allows saving reaction scheme in chemistry file formats.
- ChemAxon Marvin Document (mrv) file keeps all information including reaction conditions, catalysts, etc. MRV can also be reloaded into MarvinSketch to regenerate the reaction scheme.
- RD file format keeps the first reaction and leaves all other compounds as either reactants or products
- Mol file format takes all compounds into one mol file and acted as a mixture of seven compounds. This feature can be used to parse the individual compounds.



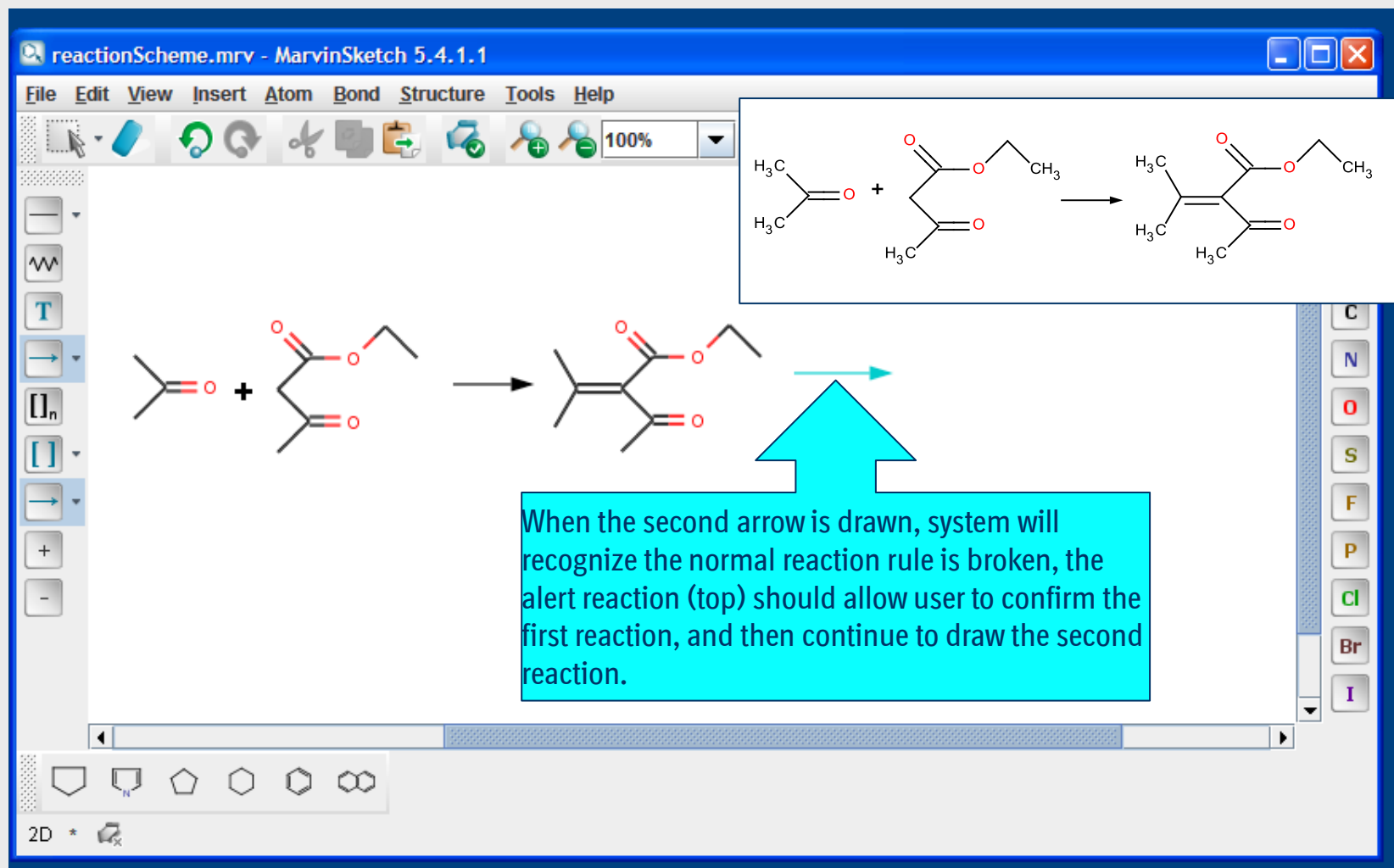
- JChem allows parsing the reaction scheme into individual compounds.
- This can only be achieved when the regular arrow, instead of reaction arrow, is used in the scheme.
- Mol file containing a mixture of all the compounds in reaction scheme can be separated using getFragments() method.
- However, the order of compounds is not necessarily consistent with the reaction scheme.
- Ideally developer should have some control of the order or at least the behavior can be understandable.

Synthesis Request(s)			
Select	Structure	Desired Amt (g)	Desired Purity (%)
1 <input type="checkbox"/>		<input type="text" value="0.0"/>	<input type="text" value="95.0"/>
2 <input type="checkbox"/>		<input type="text" value="0.0"/>	<input type="text" value="95.0"/>
3 <input type="checkbox"/>		<input type="text" value="0.0"/>	<input type="text" value="95.0"/>
4 <input checked="" type="checkbox"/>		<input type="text" value="5.0"/>	<input type="text" value="95.0"/>
5 <input checked="" type="checkbox"/>		<input type="text" value="1.0"/>	<input type="text" value="95.0"/>
6 <input checked="" type="checkbox"/>		<input type="text" value="1.0"/>	<input type="text" value="95.0"/>
7 <input checked="" type="checkbox"/>		<input type="text" value="0.1"/>	<input type="text" value="95.0"/>

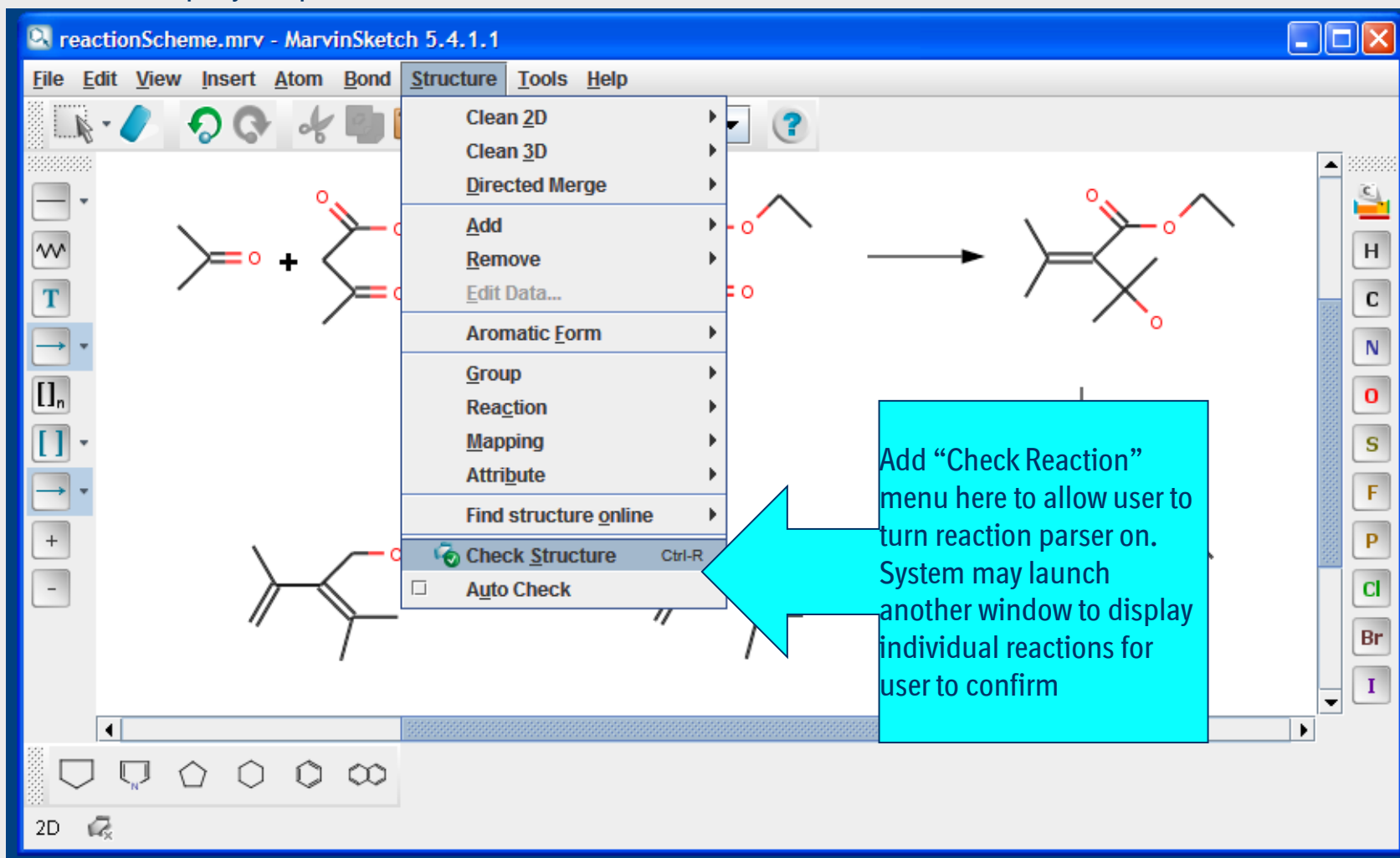
- It will be of great value to parse individual reactions, and store them in database. So far, no chemistry software can do this effectively.
- Ideally, reactions should be parsed, and ordered according to the arrow direction (up, down, right, or left).
- The generated reaction should retain the conditions, catalysts, solvents, etc.
- Not a easy task!



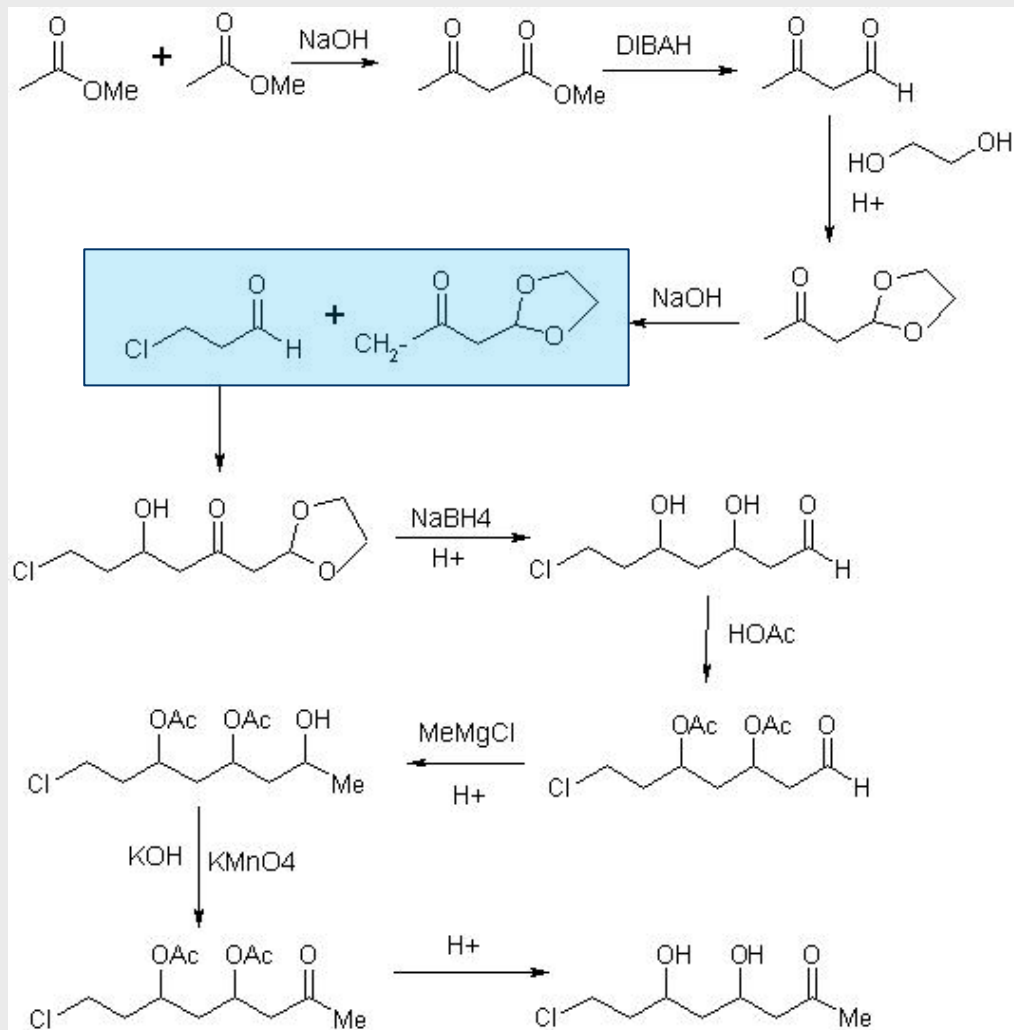
One of proposals is to allow user to confirm each individual reaction when a scheme is drawn. This feature can be switched on/off.



Another suggestion is to add menu called Check Reaction below Check Structure. This will allow user to check the reaction scheme after the complete scheme is drawn. This feature is particular useful when the step-by-step confirmation is turned off.



- Potential challenges
- No solution will be perfect
- Robust and user-friendly user interface is the key
- Step-by-step approach
- Collaboration with customers in various industries



<http://imageshack.us/photo/my-images/485/lipsynledit2le3.jpg/>

- ChemAxon JChem allows us to achieve reaction scheme capture, regeneration and compound iteration programmatically, which are difficult, if possible, to achieve with other chemistry software.
- ChemAxon MarvinSketch has various options to capture the chemistry in the reactions scheme, but cannot parse individual reactions in the scheme.
- There are a strong business need for reaction scheme and synthesis planning as well as reaction parsing.
- We proposed two possible solutions for reaction parsing, and are willing to further collaborate with ChemAxon to fulfill it.