Achieving a Synthesis
How Scientific/Technical Translation Resembles and Differs from Organic Chemical Synthesis

ATA 54th Annual Conference
San Antonio - 8 November 2013

Matthew Schlecht, PhD
Word Alchemy
wordalchemytranslation.com
mattschlecht@wordalchemytranslation.com

Achieving a Synthesis

Topics for Today

- Synthesis and Translation
  - Similarities & Differences
  - Goals, Strategies, Tools, & Execution
- How is Synthesis Done?
- How is Chemistry Coded?
- Where do Computers Come in?
- Conclusions

Achieving a Synthesis

Goals for Today

- Not for you to become organic synthesis experts!
- Get the gist of what organic synthesis is, how it's done, and why
- See parallels between what we do as translators and what organic synthesis chemists do
- Gain insights from the analogous creative processes

In other words…

This is an informative tour

Synthesis and Translation
Similarities

Organic synthesis is putting together target molecules from starting materials

Translation is putting together a target text from a starting or source text

Both involve analysis, strategy, creativity, and often painstaking experimentation

In both, quite different but equally elegant solutions can be found

Synthesis and Translation
Similarities

Entirely new solutions arise many years later

Hwæt we Gardena in geardagum þeodcyninga þrym gefrunon hu ða æþelingas ellen fremedon
(*unknown*, ca. 1000)

Hear me! We've heard of Danish heroes, ancient kings and the glory they cut for themselves, swinging mighty swords!
(*Burton Raffel*, 1963)

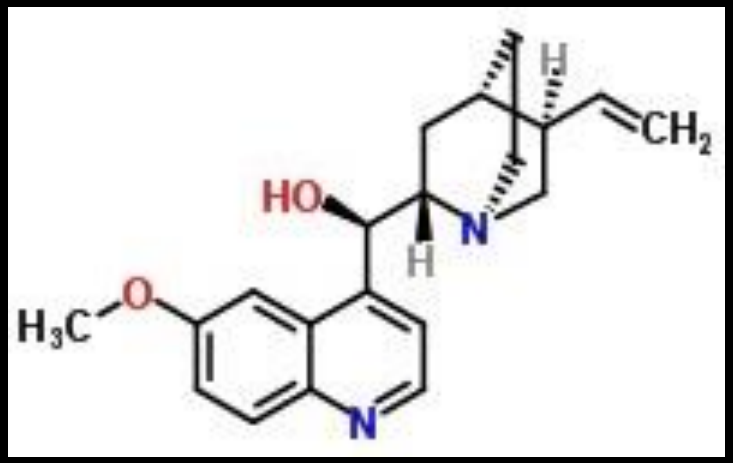
So. The Spear-Danes in days gone by and the kings who ruled them had courage and greatness. We have heard of these princes' heroic campaigns. (*Seamus Heaney*, 1999)

How we have heard of the might of the kings...
(*George Walkden*, 2013)

Synthesis and Translation
Similarities

Entirely new solutions arise many years later

Quinine



Total synthesis:

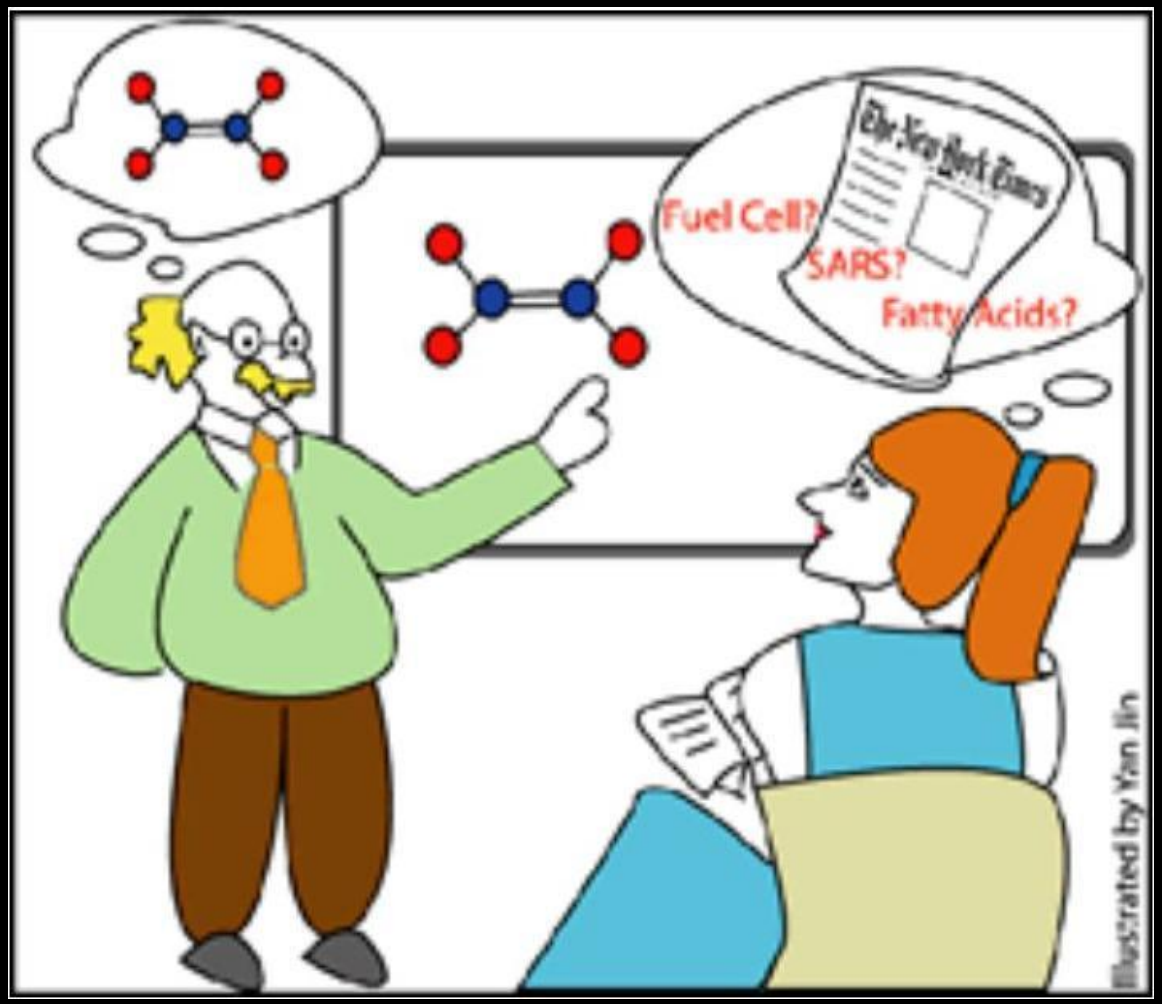
R. B. Woodward, 1945

G. Stork, 2001 (refutes Woodward)

[http://en.wikipedia.org/wiki/Quinine_total_synthesis]

Synthesis and Translation

Similarities



Synthesis and Translation
Differences

Timing

Translation: ASAP

Synthesis: months, years, decades

History

Translation: Ancient
- Karatepe inscription (bilingual, ~800 BCE)
- Behistun Inscription (trilingual, ~500 BCE)
- Rosetta Stone (trilingual, 196 BCE)

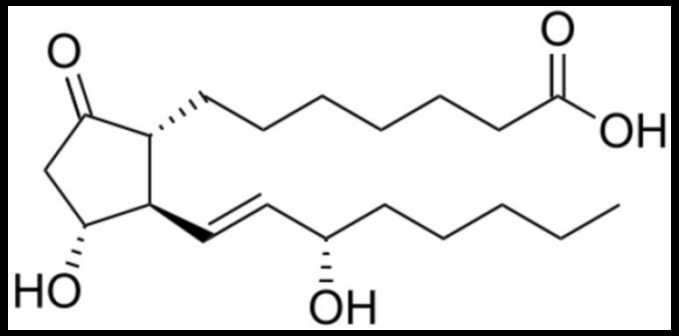
Synthesis: ~100-200 years
- Friedrich Wöhler's synthesis of urea in 1828
- Gustaf Komppa's synthesis of camphor in 1903

Reasons

Translation: usually work for hire

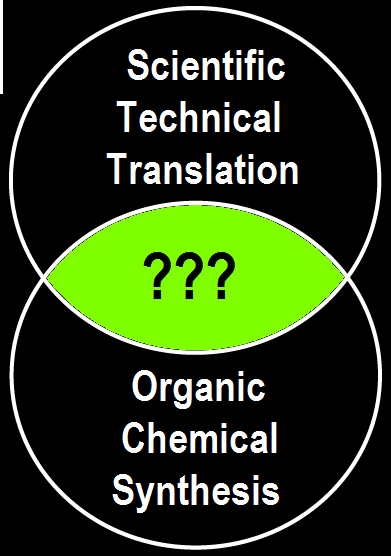
Synthesis: academic achievement, drugs, industrial process

Synthesis and Translation

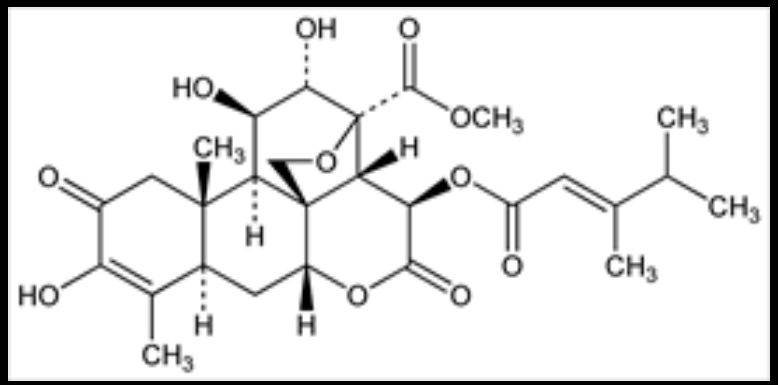


Prostaglandin E_1

Die vorliegende Erfindung betrifft Wasch- oder Reinigungsmittel mit einer speziellen α-Amylase, die durch charakteristische Sequenzvariationen gekennzeichnet ist, sowie entsprechende Wasch-oder Reinigungsverfahren und Verwendungsmöglichkeiten.

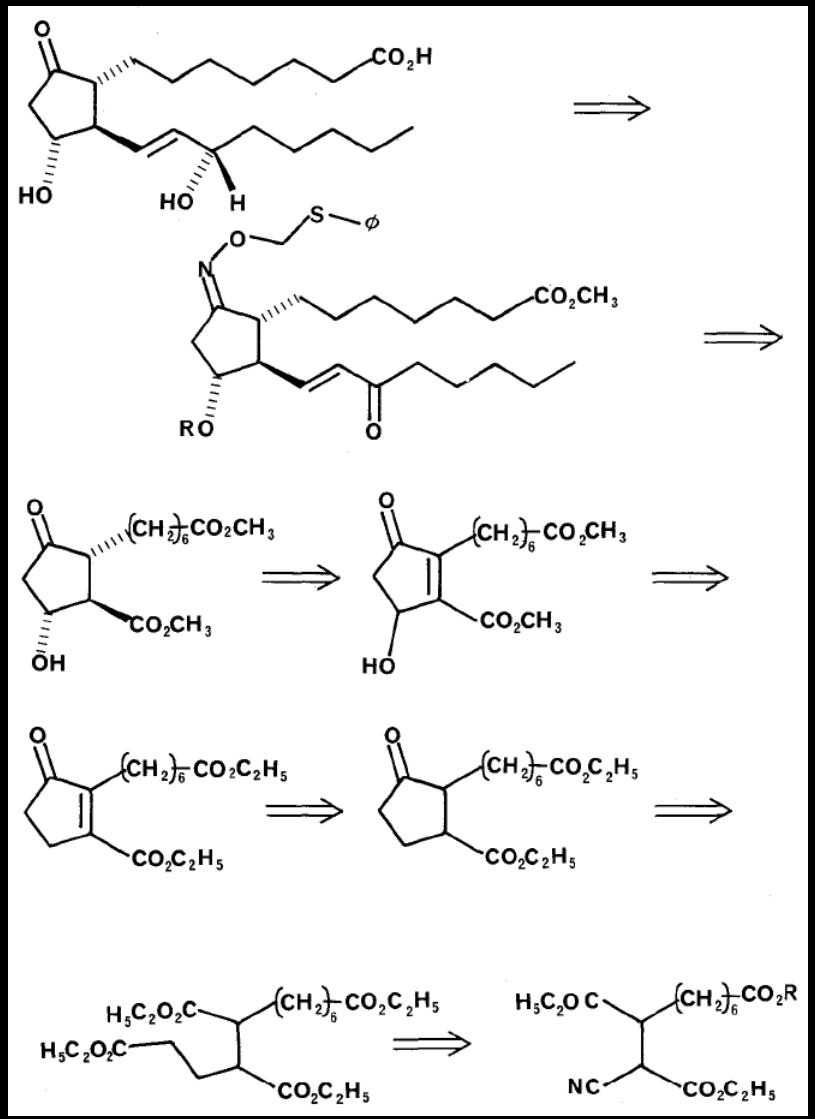


Rh陰性患者の骨盤内臓全摘術時における想定外の大量出血に対し、抗D免疫グロブリンの予防投与を行い、やむをえずRh陽性血輸血を行った症例を経験したので報告する。



Bruceantin

Synthesis and Translation



Mitra, Abhijit *The Synthesis of Prostaglandins* (New York: Wiley, 1977), p 74

Rh陰性患者の骨盤内臓全摘術時における想定外の大量出血に対し、抗D免疫グロブリンの予防投与を行い、やむをえずRh陽性血輸血を行った症例を経験したので報告する。

⇓

{0>Rh陰性患者の骨盤内臓全摘術時における想定外の大量出血に対し、抗D免疫グロブリンの予防投与を行い、やむをえずRh陽性血輸血を行った症例を経験したので報告する。<}0{>We report here on our experiences with the case of an Rh-negative patient who experienced massive hemorrhage during pelvic evisceration and was preventatively treated with prophylactic administration of anti-D immunoglobulin, then unavoidably given a transfusion of Rh-positive blood.<0}

⇓

We report here on our experiences with the case of an Rh-negative patient who experienced massive hemorrhage during pelvic evisceration and was preventatively treated with prophylactic administration of anti-D immunoglobulin, then unavoidably given a transfusion of Rh-positive blood.

Synthesis and Translation
Goals

Translation	Synthesis
The synthesis of word strings from a source language to a target language	The translation from starting materials to the target molecule

Synthesis and Translation
Strategies

Translation	Synthesis
Parse the source language word strings, then devise target language equivalents	Parse the target molecule, then devise building blocks and steps to join them together

Synthesis and Translation
Tools

Translation	Synthesis
Glossaries, corpora, reference works, CAT tools, translation memories, spell check	Starting materials, reactions, apparatus, purification means, spectroscopy

Synthesis and Translation

Execution

Translation	Synthesis
Science	Science
Art	Art
Craft	Craft

Synthesis and Translation

Is
Translation / Synthesis
an art, a science, or a craft?

Yes!

Synthesis and Translation

The **art** of elegant expression, when the beholder recognizes mastery of the medium

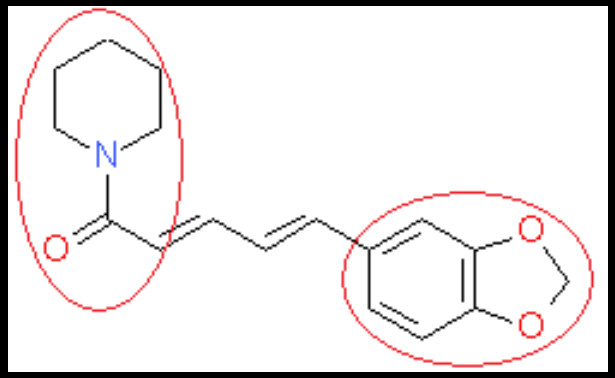
The **science** of technical know-how, using new or familiar tools to join and polish components

The **craft** of applying experience and workmanship to create a natural and useful piece of work

Synthesis and Translation

How to begin: parse for units

Piperine



Synthesis: substructures

Translation: key terms/phrases

La pipérine (ou 1-pipéroylpipéridine) est un alcaloïde au goût piquant (pseudo-chaleur).

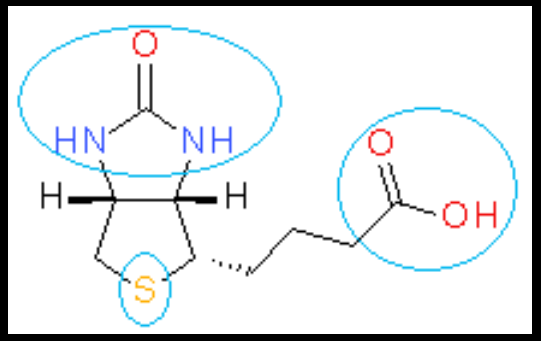
La pipérine a été découverte par Hans Christian Ørsted en 1819. On en trouve dans un champignon généralement présent dans les zones humides. C'est également le composé actif du poivre noir, blanc et du gris.

La pipérine est en partie responsable de la sensation de pseudo-chaleur lors de la consommation de poivre. Sur l'échelle de Scoville la pipérine est moins piquante (100 000 SHU) que la capsaïcine (16 000 000 SHU), le composant piquant des piments, mais plus que (6)-gingérol, le composé piquant du gingembre (60 000 SHU).

Synthesis and Translation

How to begin: parse for markers

Biotin
(ビオチン)



Synthesis: functional groups

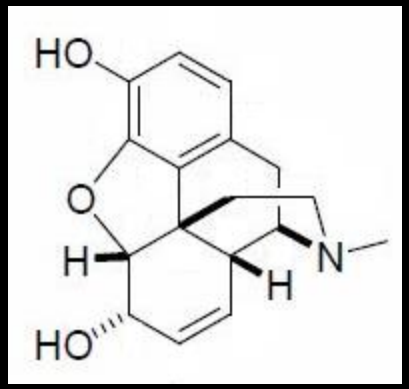
Translation: grammatical markers

ビオチンは皮膚炎予防因子として発見されたのがその始まりで、古くから皮膚病の治療に効果があるといわれてきました。現在ではアトピー性皮膚炎の治療などにビオチンが用いられています。アトピー性皮膚炎との関連では、現在様々な研究や報告がなされているところです。

Topic marker, grammatical subject, possessive, adjectival, locative/instrumental, indirect object, incomplete list conjunction

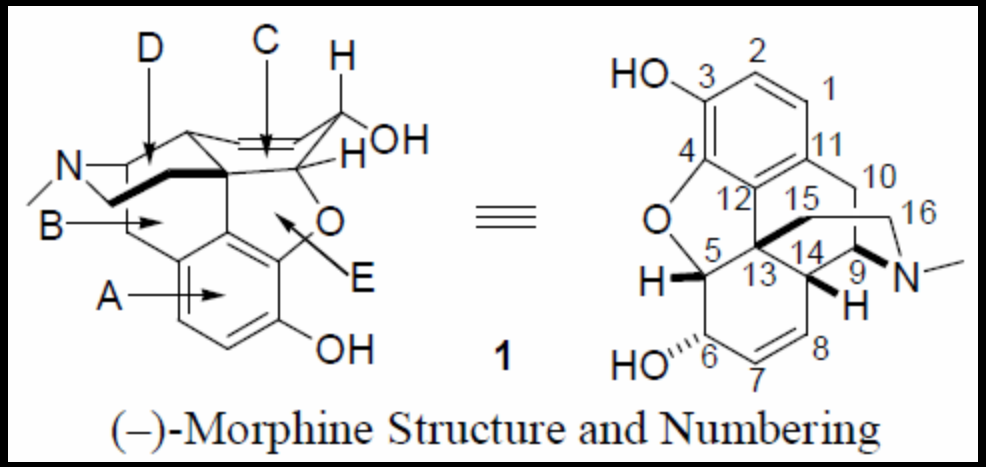
How is Synthesis Done?

Example: morphine



Why synthesize morphine?

- Prove structure?
- Industrial production?
- To make unnatural (otherwise unavailable) analogs?
- It has a challenging structure (i.e., because it's there)?

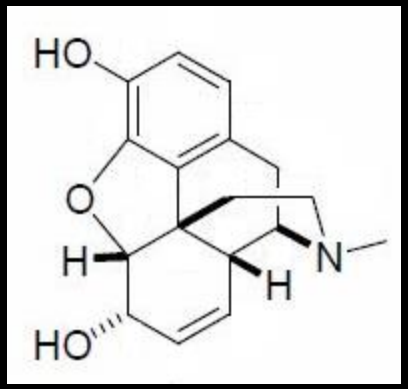


(–)-Morphine Structure and Numbering

Taber, DF; Neubert, TD; Schlecht, MF. "The Enantioselective Synthesis of Morphine". Strategies and Tactics in Organic Synthesis, Vol. 5 (Ed: Harmata, M). New York: Elsevier (2004), pp 353-389.

How is Synthesis Done?

Where to start?



What units and markers can be identified?
(*substructures and functional groups*)

How have other chemists done it?
(*dissections, strategy*)

How does Nature do it?
(*biosynthesis*)

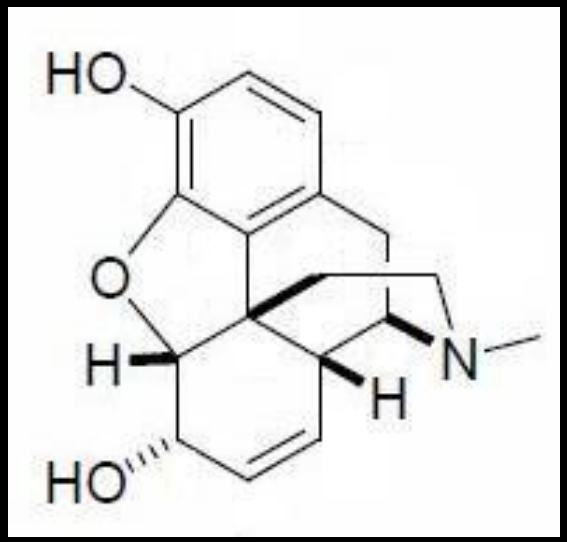
How is Synthesis Done?

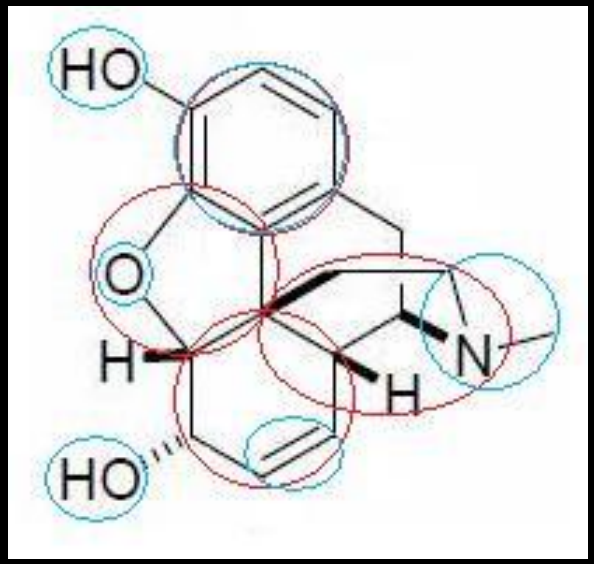
What units and markers can be identified?

Units (substructures)

&

Markers (functional groups)



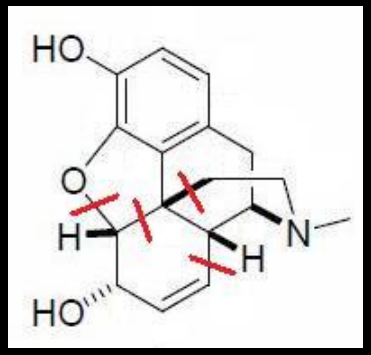


How is Synthesis Done?

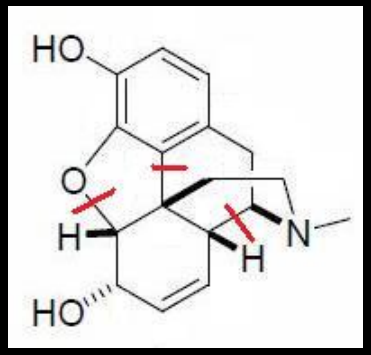
How have other chemists done it?

dissections

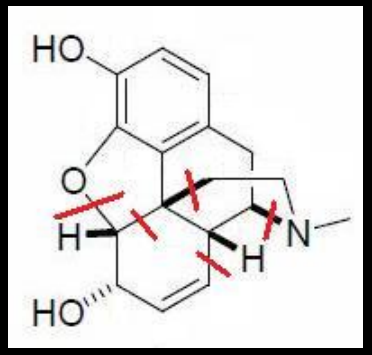
Synthesis by James White



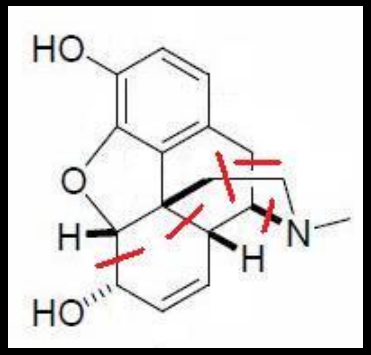
Synthesis by Kathlyn Parker



Synthesis by Marshall Gates



Synthesis by Gilbert Stork



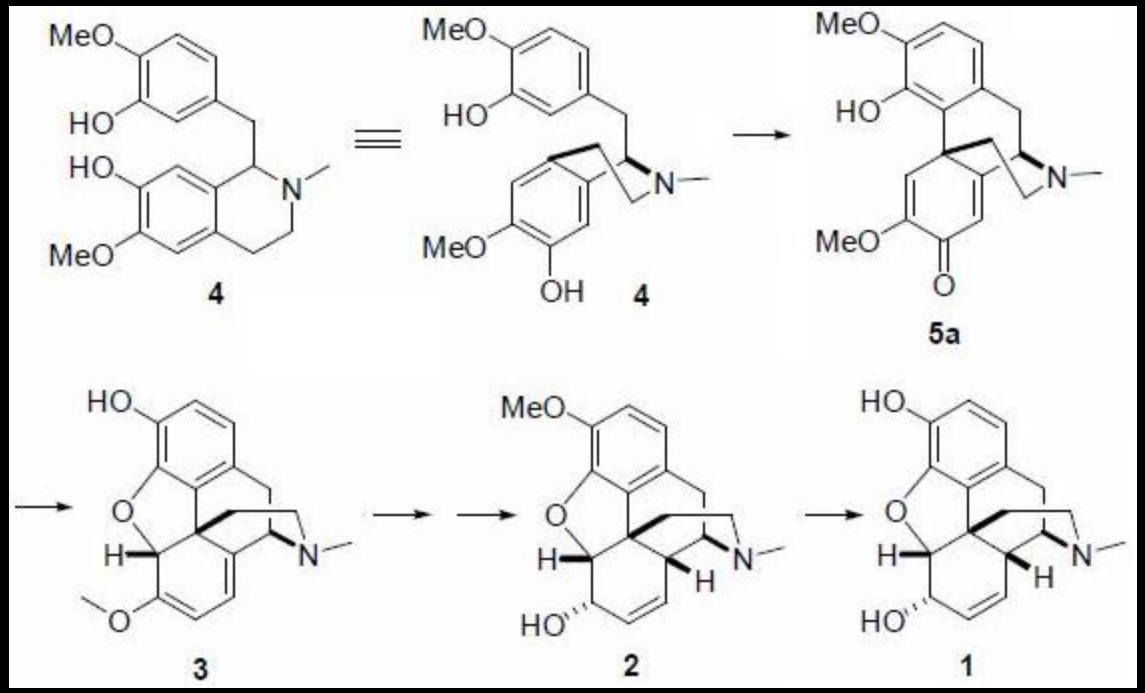
Many paths to the same destination

How is Synthesis Done?

No single correct way

How is Synthesis Done?

How does Nature do it?

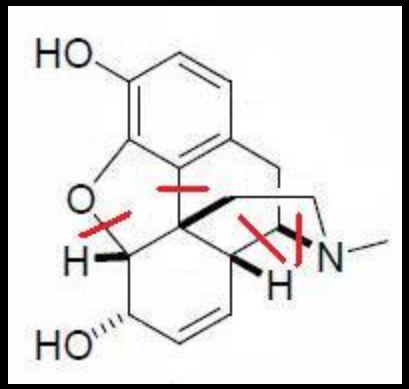


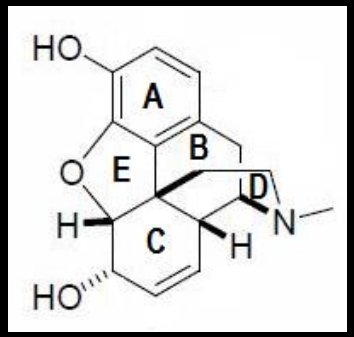
Taber, DF; Neubert, TD; Schlecht, MF. "The Enantioselective Synthesis of Morphine". Strategies and Tactics in Organic Synthesis, Vol. 5 (Ed: Harmata, M). New York: Elsevier (2004), pp 353-389.

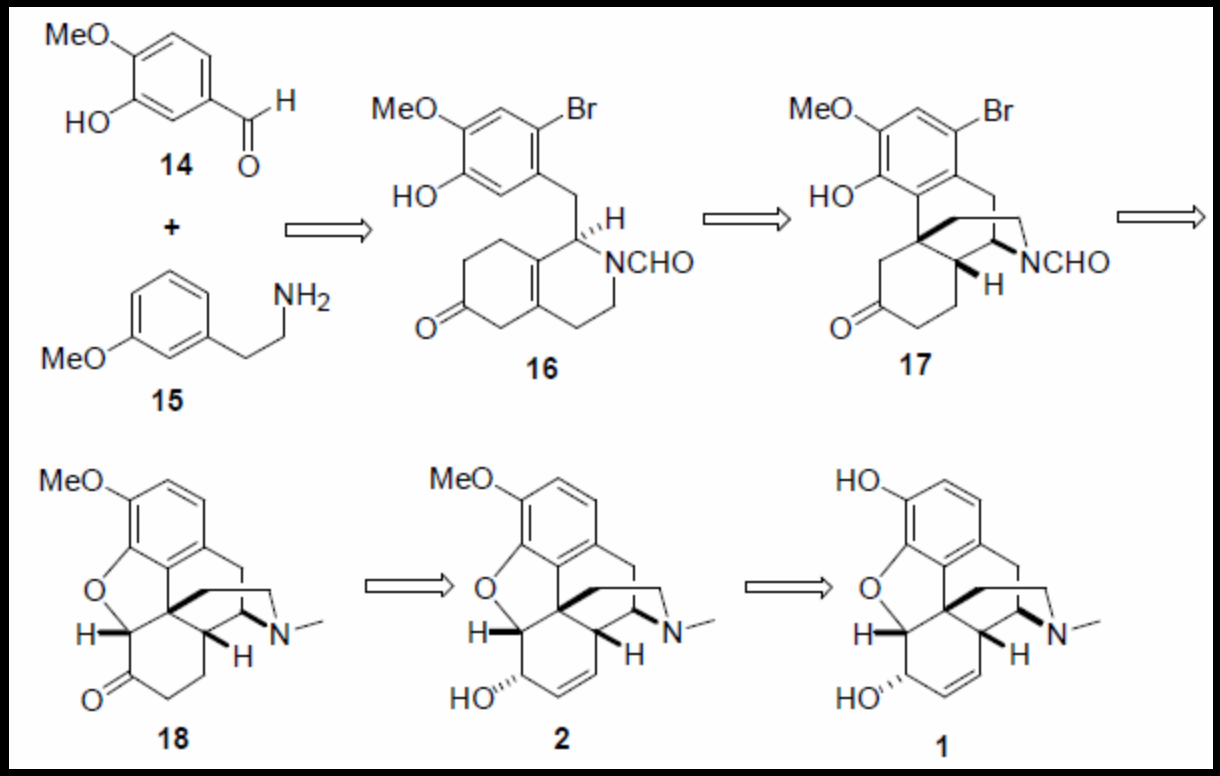
How is Synthesis Done?

Example: Synthesis by Kenner Rice

Retrosynthetic dissections







Taber, DF; Neubert, TD; Schlecht, MF. "The Enantioselective Synthesis of Morphine". Strategies and Tactics in Organic Synthesis, Vol. 5 (Ed: Harmata, M). New York: Elsevier (2004), pp 353-389.

How is Synthesis Done?

How to design the synthetic route?

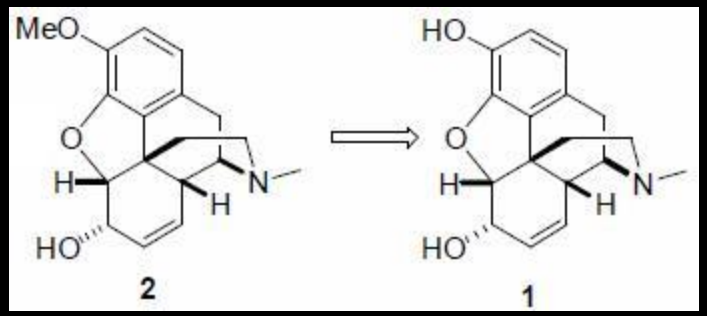
Find a way to assemble the
target molecule from starting materials

Designing a route requires extensive
knowledge of organic synthetic reactions,

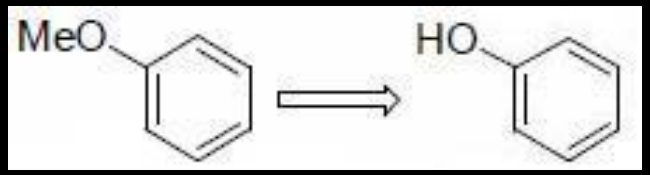
just as translation requires extensive
knowledge of terminology, phrasing, grammar

How is Synthesis Done?

The designed route requires:



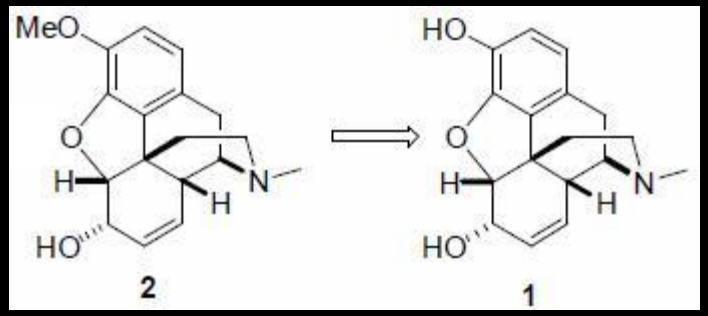
Or specifically:



Synthesis step ≈ term, phrase or translation unit

How is Synthesis Done?

One best or standard method:



use BBr_3

Analogous to a set translation:

Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß der erste Prozess...

⇩

Method according to Claim 1, characterized in that the first process…

How is Synthesis Done?

Absent a set method,
how to find the right synthesis step?

Translation analogy:

The patent **WO 2012/038571 A1** is titled

"Varilla amortiguadora porta-agujas para máquinas de tatuajes",

how to render into English the term

“varilla amortiguadora porta-agujas”

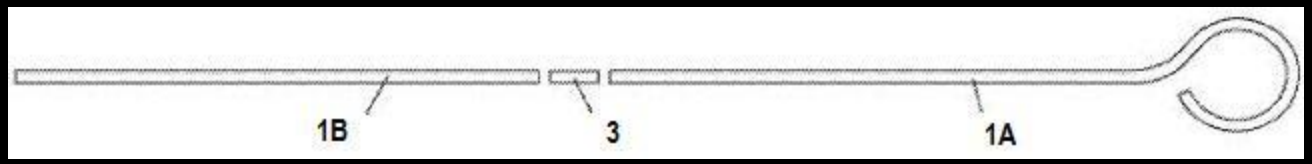
How is Synthesis Done?

Translation analogy – finding the right term:

- Look at the English version of the title, “Shock-absorbing needle bar for tattoo machines”
- Failing that, check technical glossary resources

varilla: rod, stem, bar, rib, lath, wand, link, tie, spear, link, lever, stem

- Look at the figure for any clues:

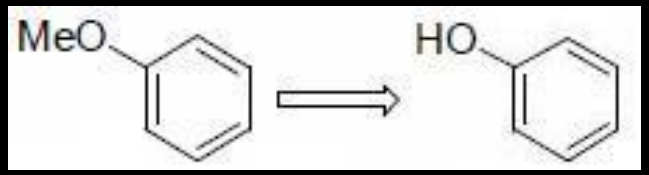


- Do context searches on the best guesses, such as googling "tattoo" with "needle ✱" where ✱ is our provisional gloss for varilla. Between "needle stem", "needle rib", and "needle bar", this latter appears to be the most widely used

How is Synthesis Done?

For synthesis, look for precedent procedures:

- Text search of chemical literature for “anisole demethylation” or “O-demethylation”
- Partial structure search for the conversion of interest:



in reaction databases (i.e., reaction corpora)

[http://www.organicworldwide.net/content/reaction-databases]

- Identify documented procedures/conditions for the conversion
- Determine compatibility in rest of the molecule (sensitivity, selectivity)
- Test experimentally.

How is Chemistry Coded?

or

How is chemistry digitized to facilitate searching and accessibility?

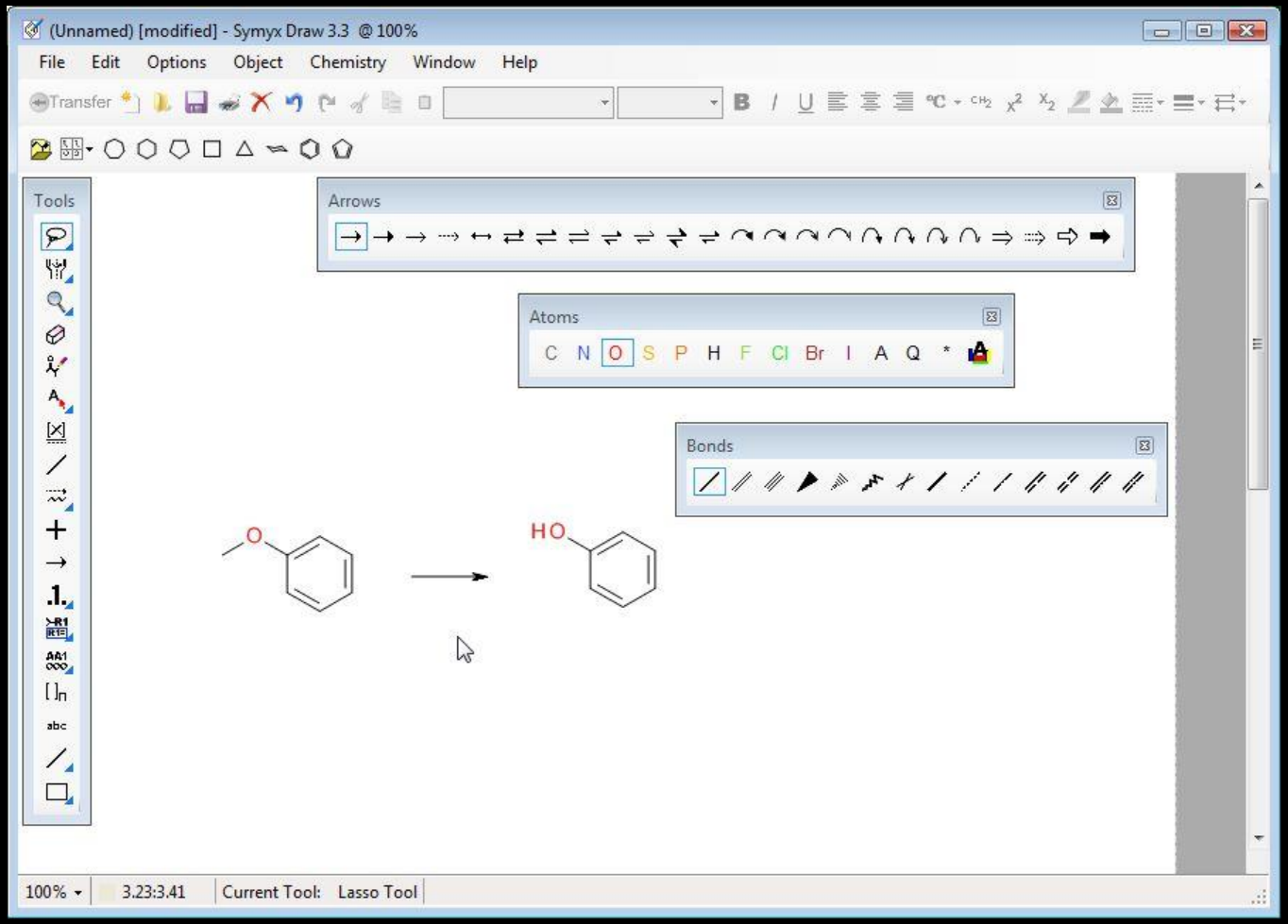
Before computers, there were written notes, note card systems, and memory.

Computers changed everything!

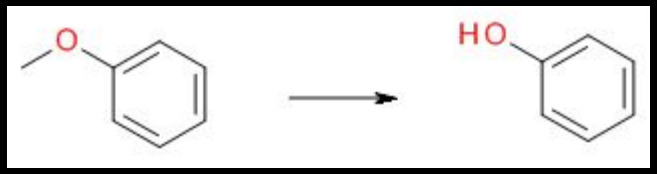
Algorithms were developed to represent molecules and reactions in digital form.

How is Chemistry Coded?

Graphical structure creation programs like **Symyx Draw**® are used
Graphic interface to create chemical reactions for a search query



How is Chemistry Coded?



Properly prepared graphic representations have chemical significance

The search query can be used to search numerous reaction databases

If too many hits, can refine the query to reflect more requirements

Sift through the results to find best possibilities to try in the lab

How is Chemistry Coded?

Digital coding of chemical structures so that the graphic has chemical significance

(similar to “dead” vs. “live” PDFs)

Many ways to do this – one way is **SMILES**

Simplified Molecular-Input Line-Entry System

Developed by chemists in an EPA lab together with academic scientists and consultants

A linear text/numerical string codes a chemical structure

How is Chemistry Coded?

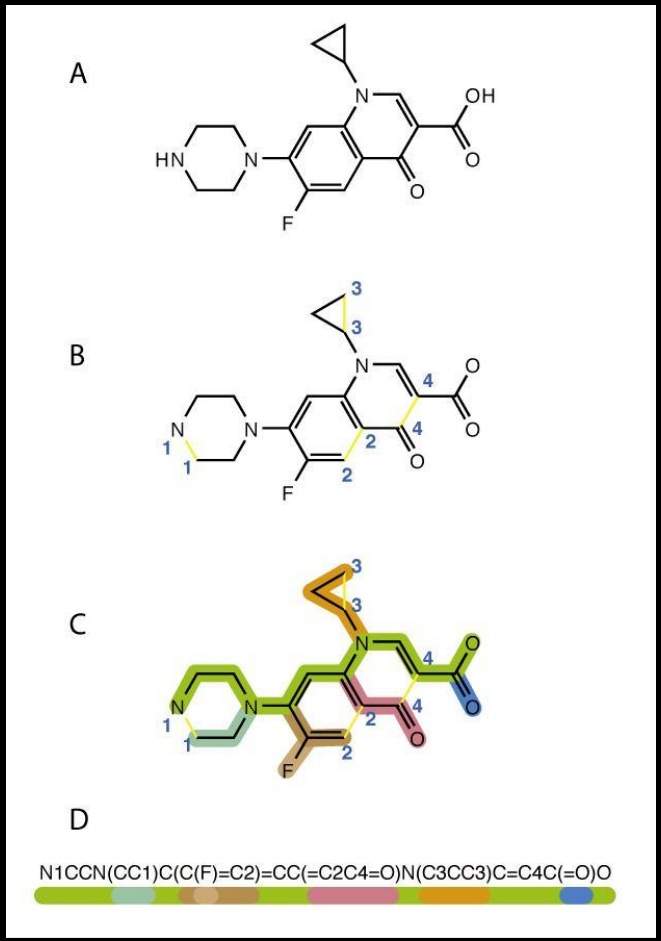
The antibiotic **ciprofloxacin**

Break the rings so that only linear strings remain

Code the breaks with numbers

Write out the SMILES code:

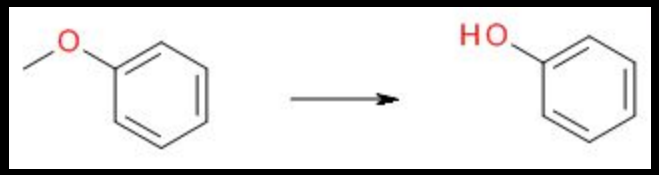
N1CCN(CC1)C(C(F)=C2)=CC(=C2C4=O)N(C3CC3)C=C4C(=O)O



http://en.wikipedia.org/wiki/Simplified_molecular-input_line-entry_system

How is Chemistry Coded?

Digitize our reaction query into SMILES:



COc1ccccc1 ➔ Oc1ccccc1

Search this coding in digitized reaction databases
Locate examples with this reaction substructure

This has analogy in the context searching that translators do for term/phrase research
Find the best fuzzy match!

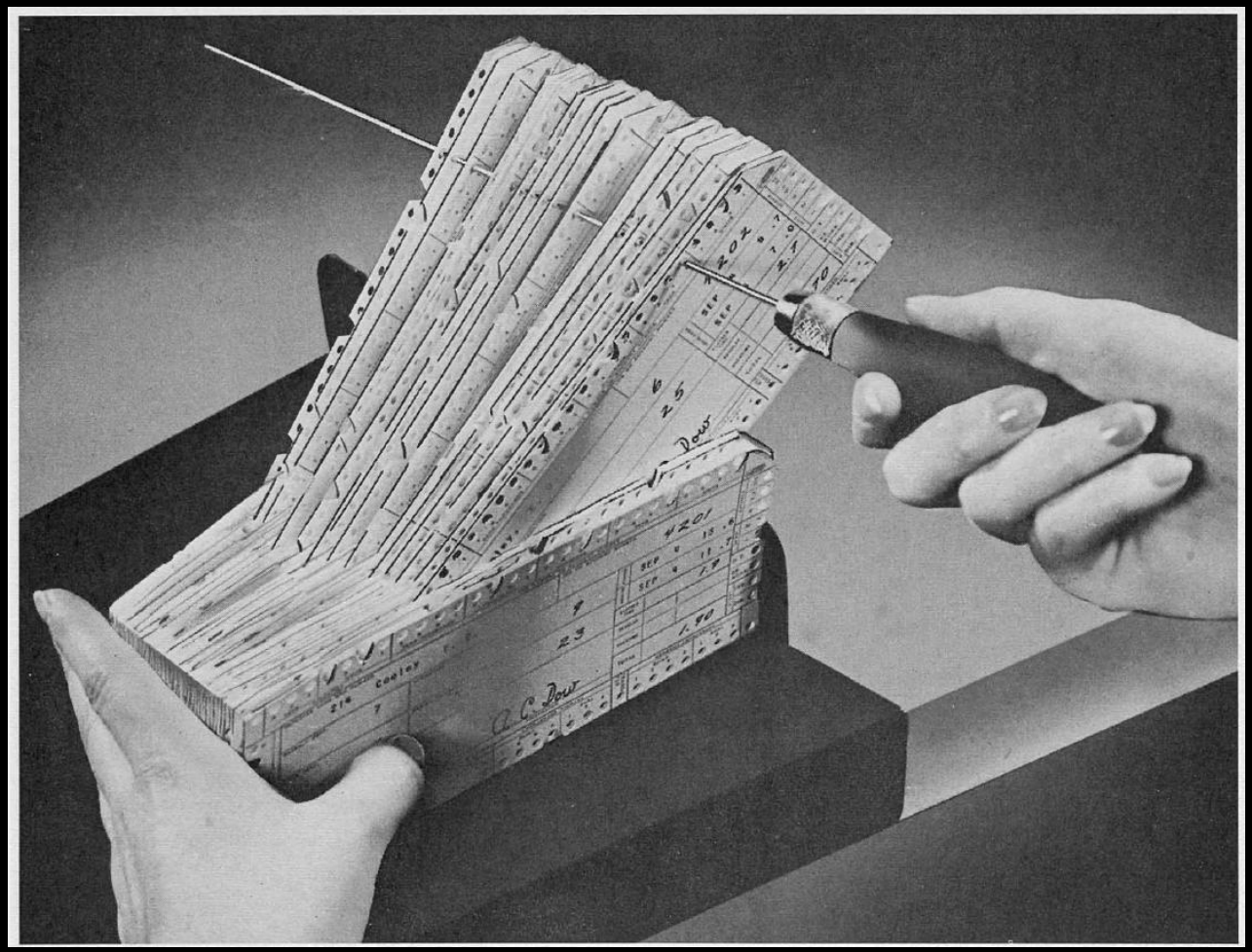
Where do Computers Come in?

Translators must manage language data: Glossaries, Corpora, Translation Memories

Synthesis chemists also must manage synthesis reaction data

Done on computers now, but wasn’t always so…

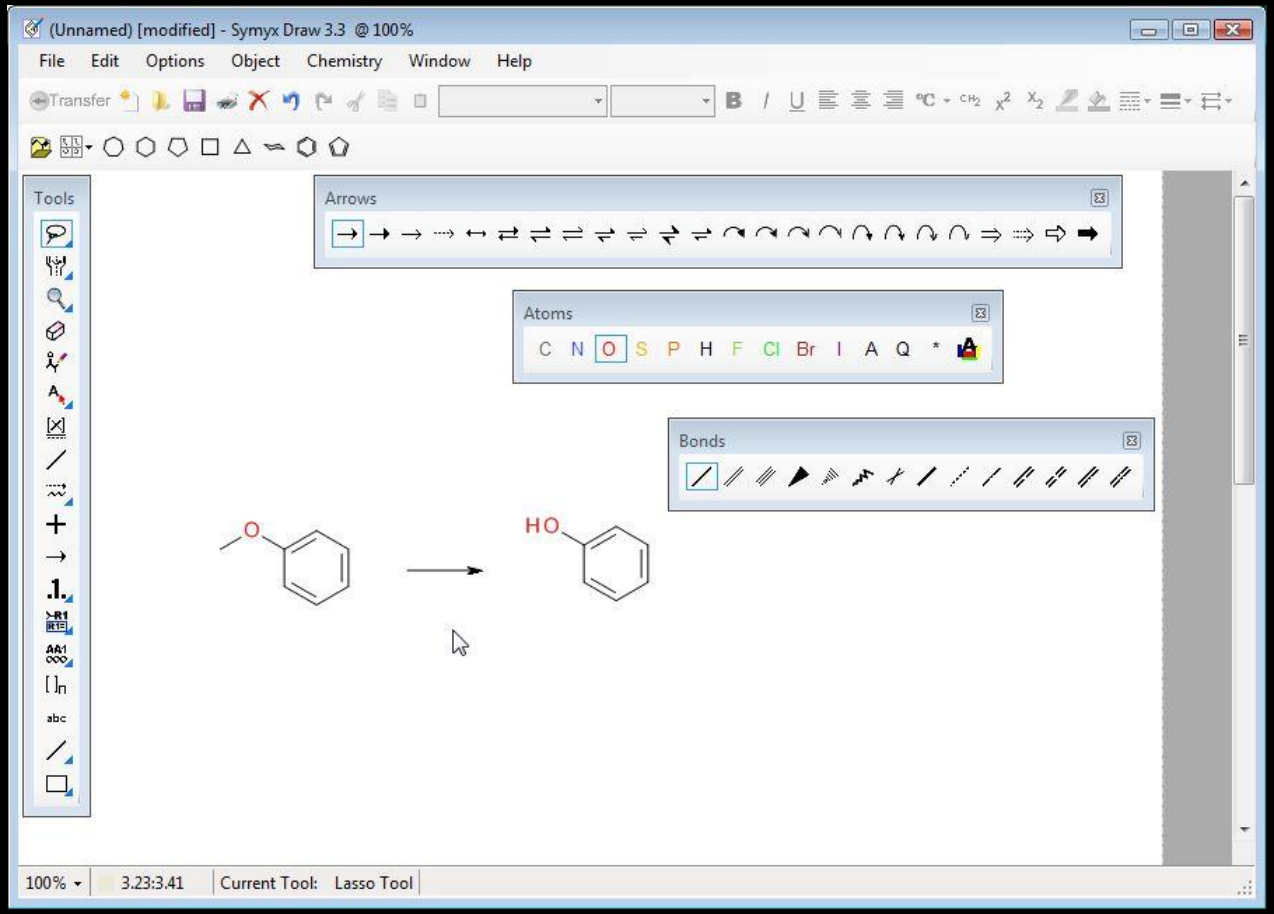
Where do Computers Come in?



State of the art, ca. 1975

Charles P. Bourne, CP. "Methods of Information Handling." New York: Wiley (1963), p 81

Where do Computers Come in?



State of the art, 2013

Symyx Draw® by Accelrys

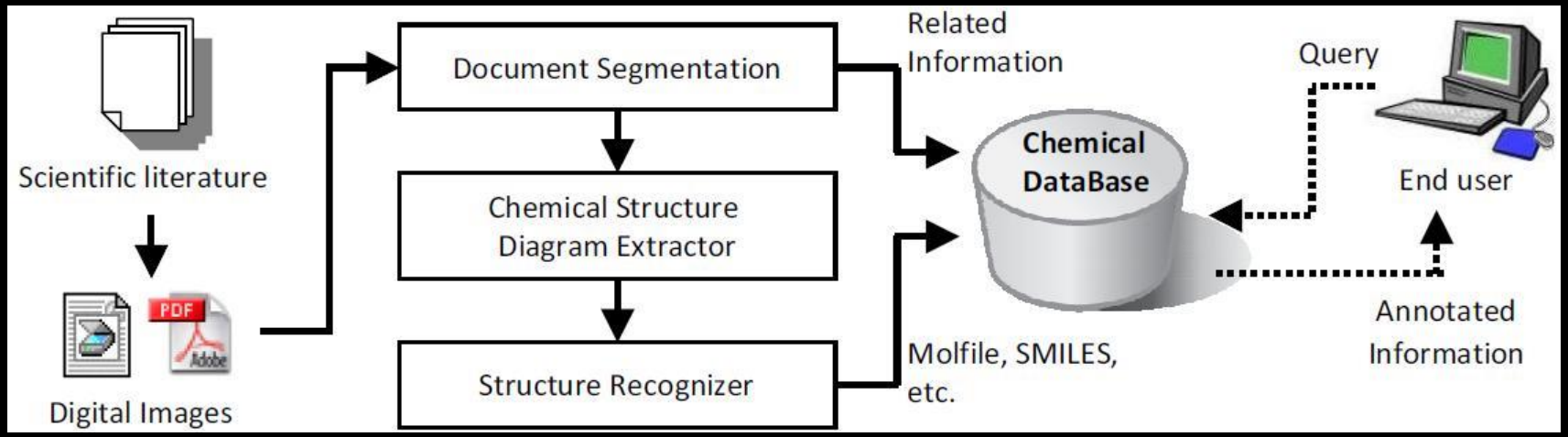
Where do Computers Come in?

Chemical literature in general, and reaction databases in particular, are where to find the data

How does the data get into databases?

Similar to the question/problem of who writes dictionaries and compiles glossaries/corpora

Where do Computers Come in?



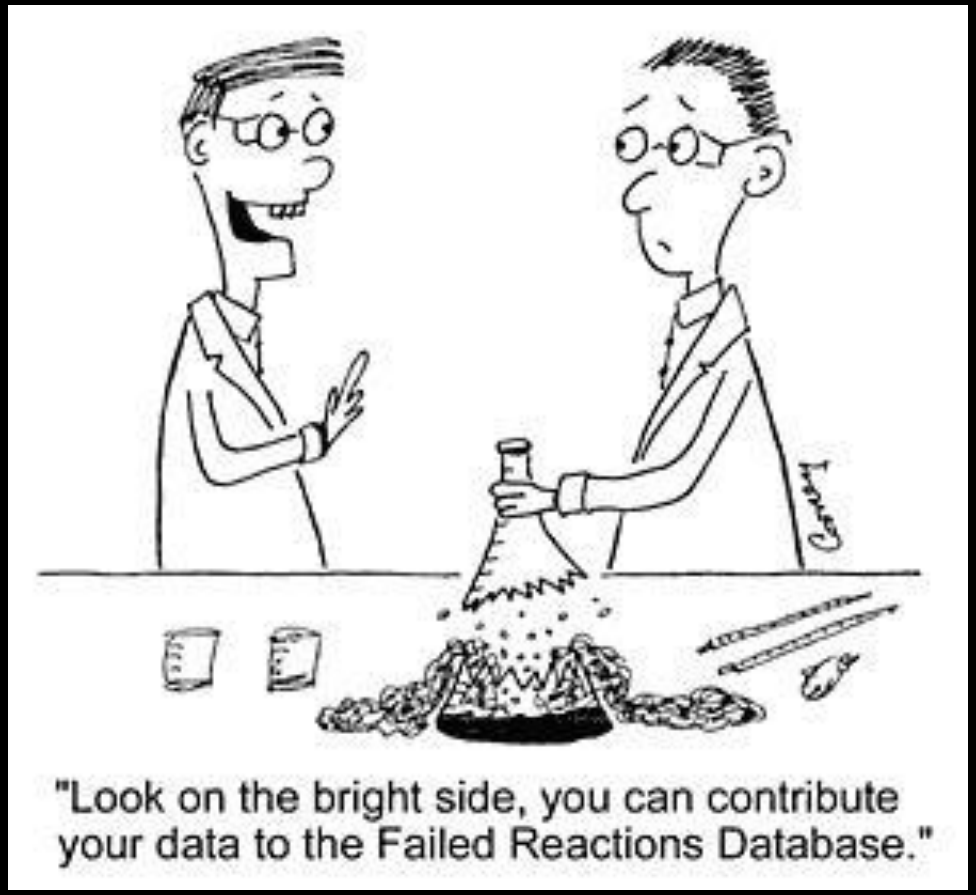
"Automated extraction of chemical structure information from digital raster images"
Park, J; Rosania, GR; Shedden, KA; Nguyen, M; Lyu, N; Saitou, K. Chem Central J **2009** 3 4
[http://journal.chemistrycentral.com/content/3/1/4]

Legacy reaction data mined from literature

Current data entered as generated

$$$

Where do Computers Come in?



Even data on failures can be welcome

Where do Computers Come in?

The next stage was:

Computer-Aided Synthesis Design

(CASD)

Computer-Assisted Organic Synthesis (CAOS)

Use computer algorithms to create synthesis routes

Sounds like a good idea, but…

Where do Computers Come in?

"...computer-assisted synthesis planning has been met with utter skepticism, even hostility, from the majority of chemists.“

Ihlenfeldt, W; Gasteiger, J.
“Computer-assisted planning of organic syntheses: the second generation of programs.”
Angew Chem Int Ed Engl **1996** 34 2613–2633.

Sound familiar?

Where do Computers Come in?

Historically…

OCSS

(Organic Chemical Simulation of Synthesis)

in 1969, followed shortly by

LHASA

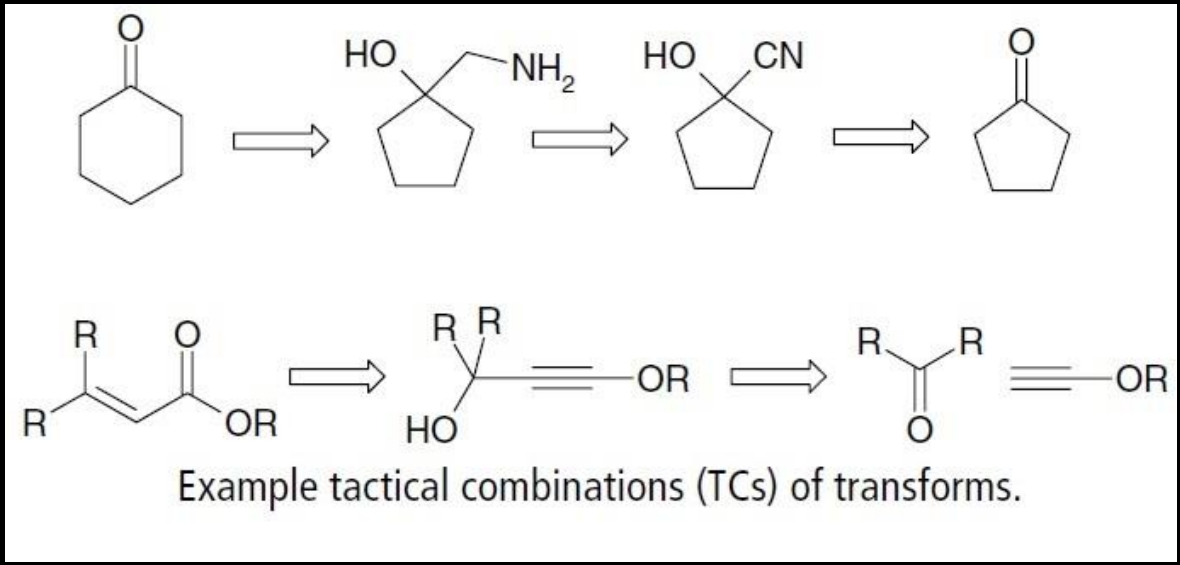
(Logic and Heuristics Applied to Synthetic Analysis)

from the E.J. Corey research group at Harvard

Development continued,
despite the frosty reception

Where do Computers Come in?

These are rule-based programs, using transforms

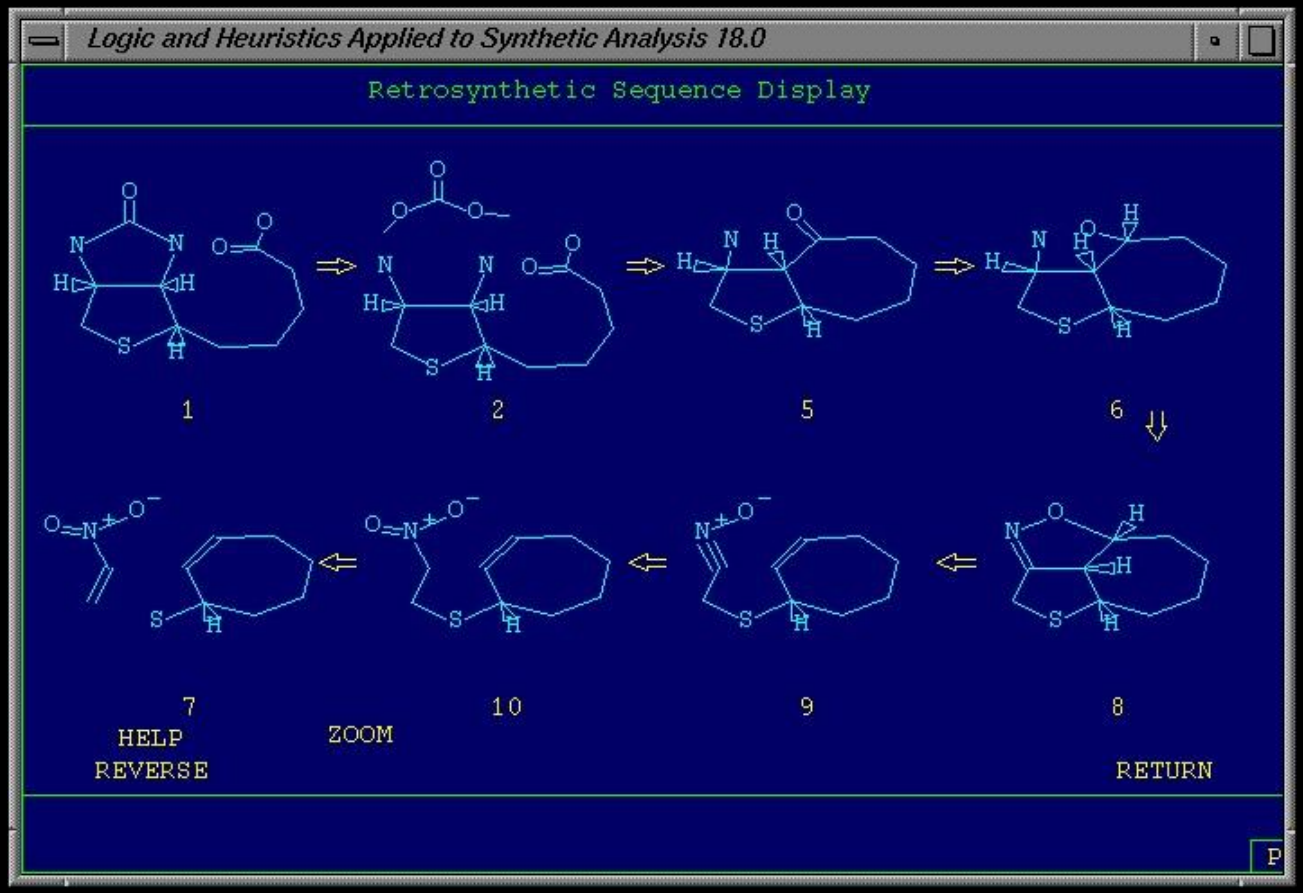


Example tactical combinations (TCs) of transforms.

Elaborately coded, reusable steps to work backwards from a target molecule

Where do Computers Come in?

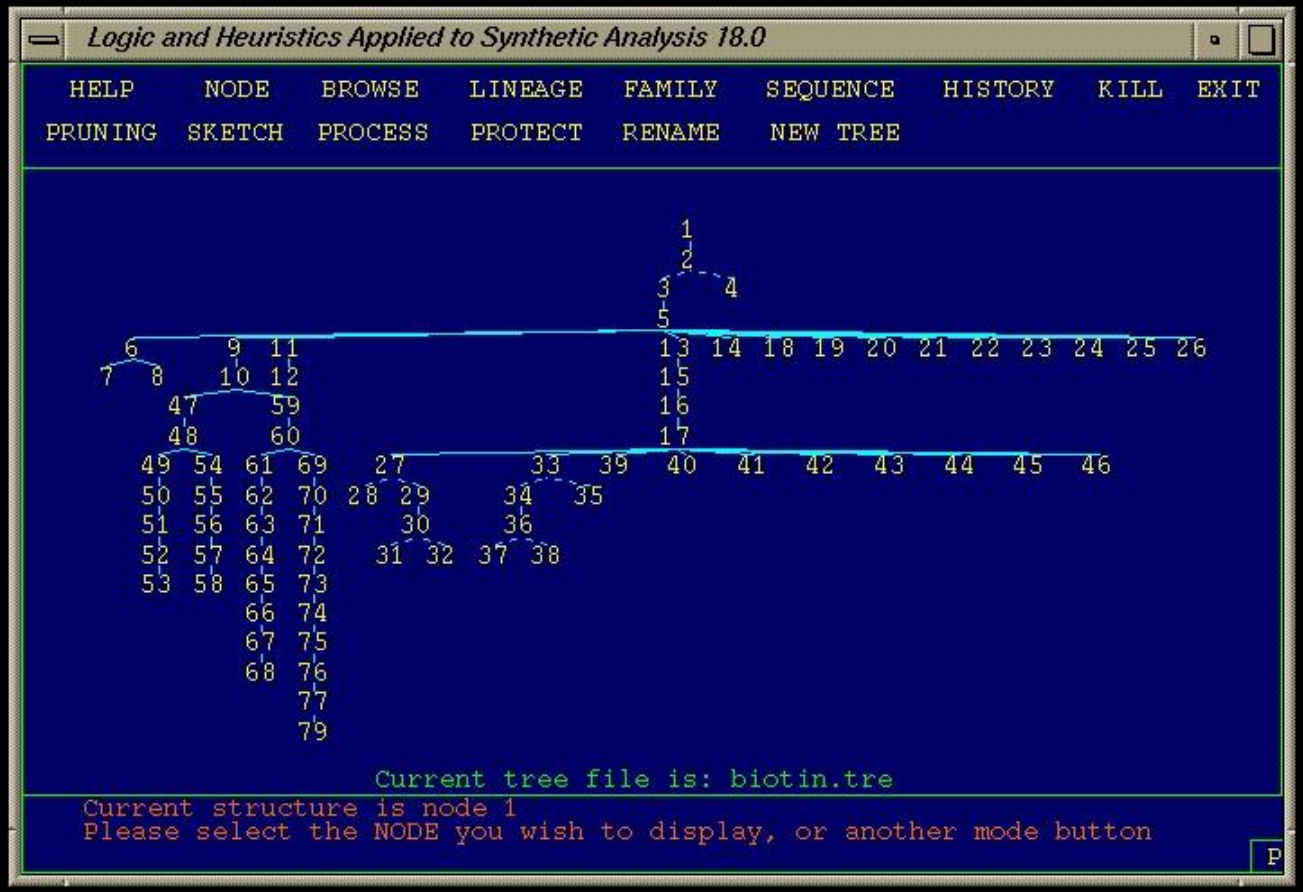
Sequence Display in the LHASA program



[http://lhasa.harvard.edu]

Where do Computers Come in?

Retrosynthetic Tree in the LHASA program



[http://lhasa.harvard.edu]

Where do Computers Come in?

Creation of new and more refined transform units was the impediment to rapid development – not scalable.

Now, automate the creation of the transform units, or generate on-the-fly based on data from reaction databases

Newer approaches are more corpora-based than rule-based

(also a current trend in MT, like GoogleTranslate)

Where do Computers Come in?

Many synthesis chemists still ask

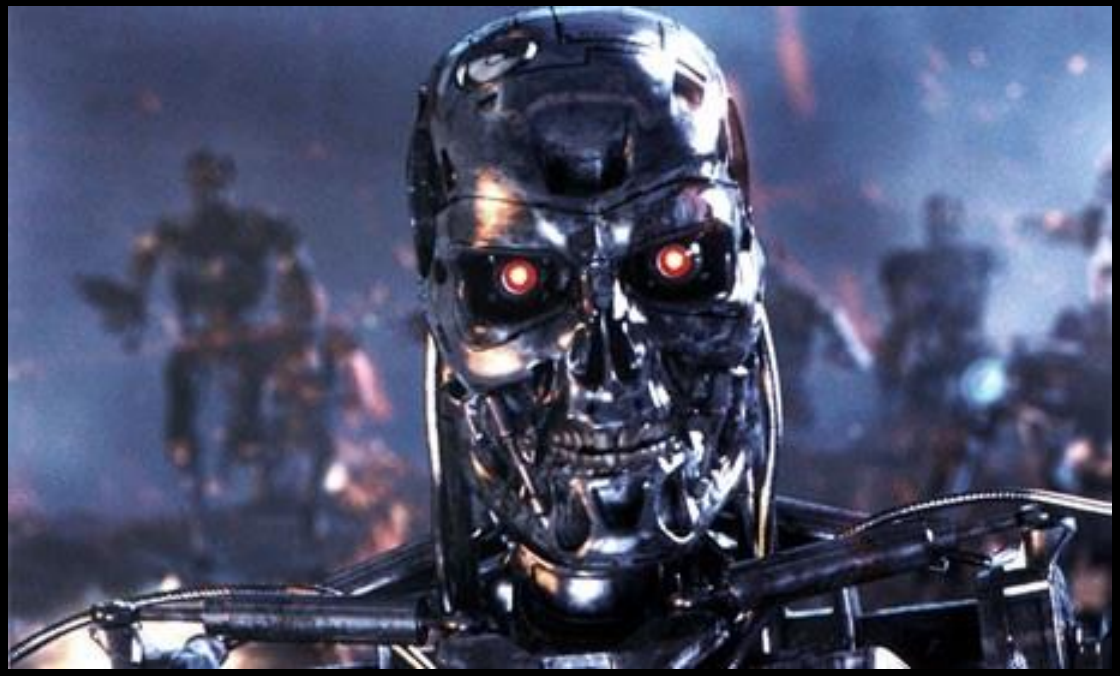
Where do computers get off?

Although they welcome reaction searching
(analogous to CAT)

Most academic synthesis chemists
still refuse to use CASD/CAOS systems
(analogous to MT)

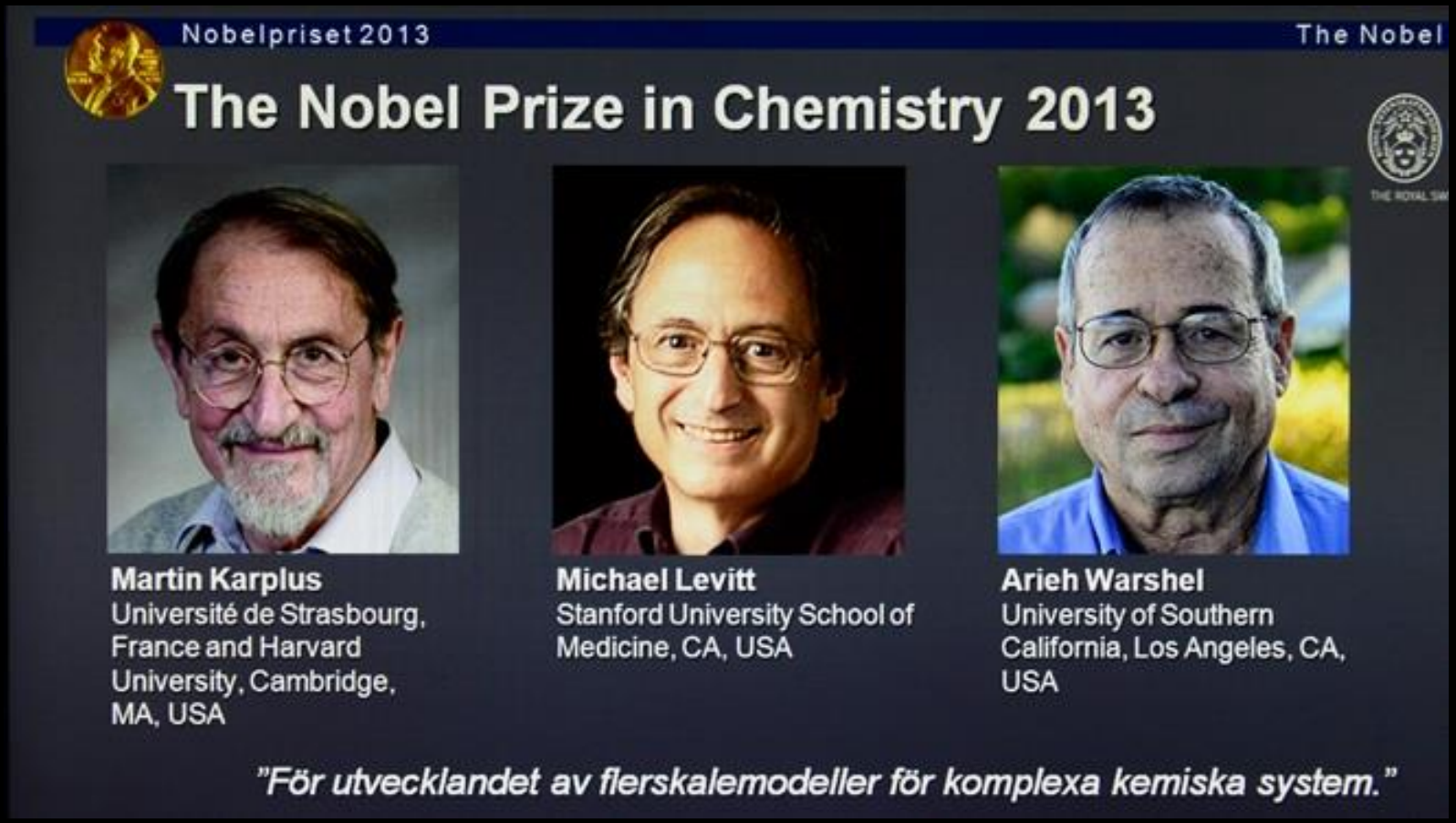
Where do Computers Come in?

Seems they're reluctant to take humans out of the loop…



Where do Computers Come in?

They’re here to stay in chemistry



“For the development of multi-scale models for complex chemical systems”

Where do Computers Come in?

Academic synthesis chemists freely use reaction database searching
(analogous to CAT)

Friendly machines to augment (not replace) the creativity of synthesis design

Conclusions

Organic synthesis and translation both do tempt amateurs, but they can't stay the course like a professional.

apologies to Graham Greene (The Third Man) and Walter White

Conclusions

Organic Synthesis and Translation

Two creative endeavors with common traits

- Best practitioners have excellent problem-solving skills
- Employ goals, strategies, tools, & execution
- Require analysis, strategy, creativity, trial, and error
- Solutions can be quite different but equally elegant
- Generally welcome computer-aided approaches
- Generally reject computer-driven approaches
- Post-editing of machine output as a “middle path”

Achieving a Synthesis

How Scientific/Technical Translation Resembles and Differs from Organic Chemical Synthesis

ATA 54th Annual Conference
San Antonio - 8 November 2013

Matthew Schlecht, PhD
Word Alchemy
wordalchemytranslation.com
mattschlecht@wordalchemytranslation.com