Open Problem:

Determine the ideal defining $\mathrm{Sec}^4(\mathbb{P}^3 \times \mathbb{P}^3 \times \mathbb{P}^3)$.

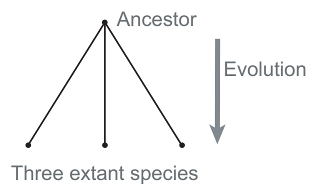
(Here $\mathrm{Sec}^4(V)$ means the closure of the union of 3-dimensional secant spaces determined by 4 points on V.)

Reward:

At the IMA workshop in March 2007, I offered a reward for a solution to this problem — an Alaskan speciality: smoked Copper river salmon. With pleasure, I will personally catch, smoke, and send Alaskan salmon to the solver(s).

Motivation:

The particular interest in $\mathrm{Sec}^4(\mathbb{P}^3 \times \mathbb{P}^3 \times \mathbb{P}^3)$ arises from molecular *phylogenetics* [AR07b], in which DNA sequences are used to infer evolutionary trees describing the descent of species from a common ancestor.



To interpret $\mathrm{Sec}^4(\mathbb{P}^3 \times \mathbb{P}^3 \times \mathbb{P}^3)$, imagine 3 currently extant species that descended from a common ancestor, as in the figure above. At a particular site in the genome of each of these species, one might observe any of the 4 nucleotides A, C, G, T. Conditioned on a particular ancestral nucleotide of A, for instance, one might expect evolution to occur in such a way that the observation of the nucleotide in each of the extant species are independent of one another. Thus for each of the 4 possible ancestral nucleotides, we have an independence model.

Such an independence model is associated with the variety $\mathbb{P}^3 \times \mathbb{P}^3 \times \mathbb{P}^3$. There are 3 independent variables (the nucleotides in the extant species), hence 3 factors in the product. Each of the variables can assume 4 values (A, C, T, G), hence its distribution is determined by a point in $\mathbb{P}^3 = \mathbb{P}^{4-1}$.

The secant variety of this Segre product represents a mixture model, obtained by mixing 4 such independence models, since the ancestral nucleotide is not known but could assume any of 4 values. The model for the 3 observed nucleotides is therefore $\mathrm{Sec}^4(\mathbb{P}^3 \times \mathbb{P}^3 \times \mathbb{P}^3)$.

In [AR07a], various theorems describe how defining polynomials for $\mathrm{Sec}^4(\mathbb{P}^3 \times \mathbb{P}^3 \times \mathbb{P}^3)$ can be used to produce polynomials for varieties associated to more general evolutionary trees, relating any number of species. Thus this particular variety is of special interest.

More generally, the varieties $\mathrm{Sec}^\kappa(\mathbb{P}^{\kappa-1} \times \mathbb{P}^{\kappa-1} \times \mathbb{P}^{\kappa-1})$ are of interest in phylogenetics for certain biologically-relevant values of κ, including $\kappa = 2, 4, 20, 61$. For $\kappa = 2$ and 3 the defining ideal for this variety is known. Thus $\kappa = 4$ is the smallest open case.

What is known?

The minimal degree of polynomials in the ideal is 5. An explicit construction of a 1728-dimensional space of quintics in the ideal was given in [AR03]. That this is the full space of quintics follows from dimension calculations in [Hag00] or [LM04]. The second of these references additionally describes this space in terms of representations.

Motivated by questions in computational complexity, Strassen [Str83] constructed a polynomial of degree 9 vanishing on $\mathrm{Sec}^4(\mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^2)$. Composed with linear maps from $\mathbb{P}^3$ to $\mathbb{P}^2$, Strassen's polynomial produces many degree 9 polynomials vanishing on $\mathrm{Sec}^4(\mathbb{P}^3 \times \mathbb{P}^3 \times \mathbb{P}^3)$, which are not in the ideal generated by the quintics. (More generally, relations between the ideals of secant varieties of Segre products of projective spaces of different dimensions are developed in [AR07a].)

A reasonable guess is that the known degree 5 and 9 polynomials generate the full ideal. Unfortunately, it is not even known whether these polynomials set-theoretically define the variety.

References

[AR03] Elizabeth S. Allman and John A. Rhodes. Phylogenetic invariants for the general Markov model of sequence mutation. *Math. Biosci.*, 186:113–144, 2003.

[AR07a] Elizabeth S. Allman and John A. Rhodes. Phylogenetic ideals and varieties for the general Markov model. *Adv. in Appl. Math.*, 2007. To appear, `arXiv:math.AG/0410604`.

[AR07b] Elizabeth S. Allman and John A. Rhodes. Phylogenetics. In Reinhard Laubenbacher, editor, *Modeling and Simulation of Biological Networks*, Proc. Sympos. Appl. Math. Amer. Math. Soc., Providence, RI, 2007.

[Hag00] Thomas R. Hagedorn. A combinatorial approach to determining phylogenetic invariants for the general model, 2000. Technical report, Centre de recherches mathmatiques.

[LM04] J. M. Landsberg and L. Manivel. On the ideals of secant varieties of Segre varieties. *Found. Comput. Math.*, 4(4):397–422, 2004.

[Str83] V. Strassen. Rank and optimal computation of generic tensors. *Linear Algebra Appl.*, 52/53:645–685, 1983.